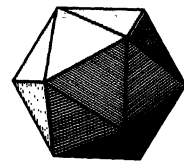


The American Mathematical Monthly



Volume 100, Number 1 / JANUARY 1993



Benjamin Franklin Finkel

AN OFFICIAL PUBLICATION OF THE MATHEMATICAL ASSOCIATION OF AMERICA

NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK
PETER BORWEIN
RICHARD BUMBY
DENNIS DETURCK
UNDERWOOD DUDLEY
JOHN DUNCAN
JOAN FERRINI-MUNDY
JOSEPH GALLIAN
STEVEN GALOVICH
RICHARD GUY
DARRELL HAILE
PAUL HALMOS
CATHERINE MCGEOCH
RICHARD NOWAKOWSKI
LEE RUBEL
LYNN STEEN
STAN WAGON
DOUGLAS WEST
HERBERT WILF

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

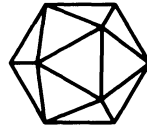
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International, Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 1 / JANUARY 1993
(ISSN 0002-9890)



Contents

ARTICLES

- The Seventy-Fifth Anniversary Celebration / G. BAILEY PRICE 4
An Axiomatic Approach to the Integral / LEONARD GILLMAN 16
Versatile Coins / ISTVÁN SZALKAI and DAN VELLEMAN 26
Two-Year Magazine Subscription Rates / UNDERWOOD DUDLEY 34
On Seeing Progressions of Constant Cross Ratio / R. J. DUFFIN 38
100 Years of *Monthly* Editors / JOHN EWING 48
Quotients of Primes / DAVID HOBBY and D. M. SILBERGER 50
Pathological Functions for Newton's Method / GEORGE C. DONOVAN,
ARNOLD R. MILLER, and TIMOTHY J. MORELAND 53
-

FEATURES

Finkel Letter 2

COMMENTS 3

NOTES 59

THE AUTHORS 67

LETTERS 69

UNSOLVED PROBLEMS

When Does a Polynomial Over a Finite Field Permute the Elements
of Field?, II / RUDOLF LIDL and GARY L. MULLEN 71

PROBLEMS AND SOLUTIONS 75

REVIEWS

Iteration of Rational Functions by Alan F. Beardon /
ROBERT L. DEVANEY 90

TELEGRAPHIC REVIEWS 94

The Seventy-Fifth Anniversary Celebration

G. Baley Price

The record [1] shows that 104 men and women, meeting on December 30 and 31, 1915, in room 101 of Page Hall on the campus of Ohio State University, formed a new organization which they named the Mathematical Association of America. They elected officers as follows: President, E. R. Hedrick of the University of Missouri; Vice-Presidents, E. V. Huntington and G. A. Miller of Harvard and Illinois; and Secretary-Treasurer, W. D. Cairns of Oberlin. They also elected an Executive Council whose twelve members were widely representative of important centers for the study of mathematics from the East Coast to the West Coast. This first Executive Council included J. W. Young, R. C. Archibald, Oswald Veblen, E. H. Moore, Florian Cajori, D. N. Lehmer, and other leaders of American mathematics. Finally, the Executive Council appointed a Committee on Publications consisting of H. E. Slaughter, R. D. Carmichael, and W. H. Bussey, with Slaughter as the Managing Editor of *The American Mathematical Monthly*.

Today we salute these pioneers whose vision and wisdom established the Mathematical Association of America; today we celebrate the seventy-fifth anniversary of the organization which they bequeathed to us.

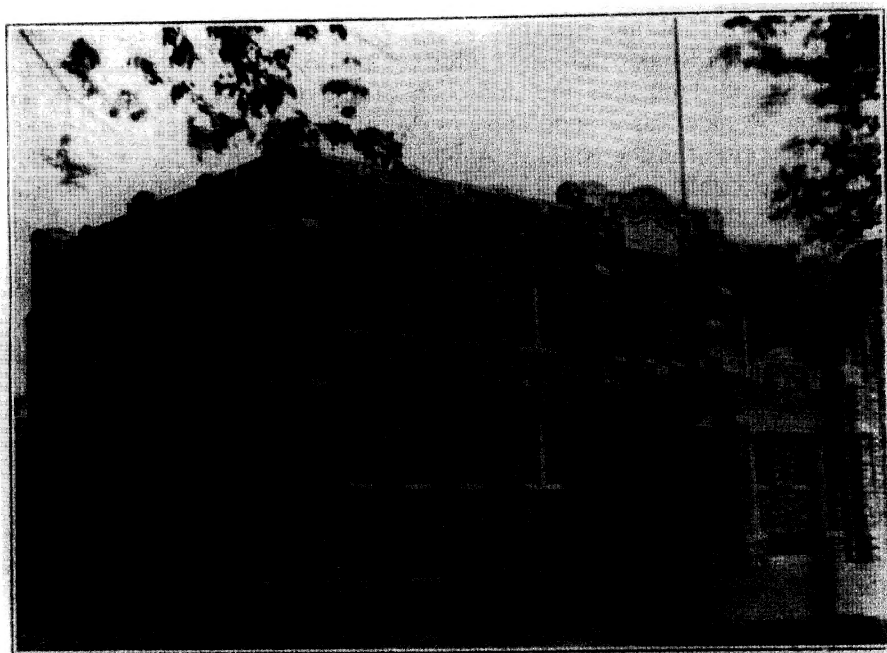
But there was an earlier beginning which must be recognized; let me first describe the setting for this event. In northwest Missouri, about 75 miles north of Kansas City and 35 miles east of Saint Joseph, there is in the county of Caldwell a small town named Kidder. It is a little town: in 1900 its population was 357, and in 1910 it was only 306. Kidder is shown on the map of Missouri in the Rand McNally Road Atlas today, but without help it is difficult to locate: it is so small that it is not included in the index of towns and cities. The neighborhood of Kidder today is a pleasant and prosperous agricultural region, but before, during, and after the Civil War it suffered much violence and terrorism [7]. But peace and stability came eventually, and in 1884 the Congregational Church revived an earlier school in Kidder and called it Kidder Institute [5, pp. 294–296], [6, pp. 553–558].

In 1892 Benjamin Franklin Finkel [8], [9] joined the faculty of Kidder Institute as a mathematics teacher. Finkel was born in Ohio in 1865, he received a B.S. degree from Ohio Northern University in 1888, and, before he arrived in Kidder, he had been a teacher, a mathematics instructor, a principal, and a superintendent in schools and academies in Ohio and Tennessee. His experiences as a teacher had created in him a desire which he described as follows:

Knowing that the status of the mathematical teaching in our high schools and academies was very deplorable and even worse in the rural schools, I had the ambition to publish a journal devoted solely to mathematics and suitable to the needs of teachers of mathematics in these schools.



KIDDER INSTITUTE.



ADMINISTRATION BUILDING, KIDDER INSTITUTE

Thus in 1893 Finkel planned the publication of a journal. He was only 28 years old and he had never been to graduate school, but he had published a book [10]. Since the journal would be his personal property, he was not obliged to seek the approval of any committee, governing body, or granting agency, but neither did he have the support of a sponsor for his journal. Mr. E. J. Chubbuck, editor of the local newspaper in Kidder, agreed to print, bind, and mail the new journal. J. M. Colaw, Principal of the high school in Monterey, Virginia, accepted Finkel's invitation to be the co-editor of his journal. Finkel wrote later [9, p. 309]: "I decided to call the new publication *The American Mathematical Monthly*, a most ambitious title, as my friend Dr. E. H. Moore afterward told me." The first number of volume 1 appeared in January, 1894.

The Congregational Church supported at that time not only Kidder Institute but also Drury College [11, pp. 813–819] in Springfield in southwest Missouri; Springfield is near the peaceful Ozark Mountain region made famous by Harold Bell Wright's book entitled *The Shepherd of the Hills* [12]. In June, 1895, through the influence of Dr. Henry Hopkins, pastor of the First Congregational Church in Kansas City, Missouri, a member of the Board of Trustees of Kidder Institute and of Drury College, Finkel was elected to the professorship of mathematics and physics at Drury College [9, p. 313]. Finkel's new position forced him to move the publication of the *Monthly* to Springfield; he was fortunate to find another local printer, Mr. Dixon, to continue the work which had been begun by Mr. Chubbuck in Kidder.

The circumstances surrounding Finkel's founding of *The American Mathematical Monthly* emphasize that it was an audacious undertaking. Others far more able than Finkel and in the leading centers of learning in the nation had failed in their efforts to establish journals. For example, Benjamin Peirce and Joseph Lovering founded a quarterly entitled *Cambridge Miscellany of Mathematics, Physics, and Astronomy* at Harvard [13], but an insufficient number of subscribers forced them to discontinue the journal after publishing only four numbers during 1842–1843. Also, John Daniel Runkle published three volumes of his *Mathematical Monthly* in 1858, 1859, and 1860, but the approach of the Civil War prevented further publication. Runkle was a pupil and protégé of Benjamin Peirce, and in 1857 he was senior assistant in the Nautical Almanac Office in Cambridge, Massachusetts. Runkle was an able person: he was President of the Massachusetts Institute of Technology from 1870 to 1878 and he was head of its Department of Mathematics until his death in 1902 [14, pp. 163–166], [15]. Simon Newcomb stated that Runkle's *Mathematical Monthly* had a beneficial influence on the development of mathematics in America [16] although only three volumes were published.

Finkel also failed—at least he failed in his stated purpose, which was to publish a journal which would be widely read by high school teachers and which would improve the teaching of mathematics in the high schools. In commenting on his campaign for subscribers for the *Monthly*, Finkel wrote as follows [9, p. 309]:

The first person to respond to my solicitation was the distinguished, scholarly, and eminently successful superintendent of Kansas City schools, Professor J. M. Greenwood. He enclosed his check for \$2.00 in payment of his subscription for one year—the first money received with which to found the *Monthly*—and he assured me that he would call the attention of all his mathematics teachers to the new venture.

Although the Superintendent of Schools in Kansas City subscribed to the *Monthly*, few other high school teachers did. Finkel wrote later that during his 19 years as Editor, not more than a dozen high school teachers were on the subscription list at any one time [9, p. 310]. Shunned by the high school teachers it was designed to serve, “the *Monthly* soon adapted itself to the needs of the field of collegiate mathematics” [9, p. 310]. But *survive* the *Monthly* did—volume 97 is being published during this current year 1990. It is now in order to explain how the *Monthly* was transformed, how it was saved, and how it became the official journal of the Mathematical Association of America and one of the important mathematics journals in America.

Mathematics developed rapidly in America in the period between Runkle’s last volume in 1860 and Finkel’s first volume in 1894. In particular, the New York Mathematical Society was established in 1888 and renamed the American Mathematical Society in 1894. The title page of issue number 1 of volume 1 of the *Monthly*, dated January, 1894, shows that both B. F. Finkel and his co-editor, J. M. Colaw, were members of the New York Mathematical Society. The first volume of the *Bulletin of the New York Mathematical Society* is dated October 1891 to July 1892 with pages 1 to 241. This volume contains also a List of Members, dated June 1892 and with pages 1 to 22. On page 8 of this List of Members we find the following:

Benjamin F. Finkel, M.Sc.
Superintendent of Schools, North Lewisburg,
Champaign County, Ohio.
Date of Admission: June, 1891.

E. H. Moore was a member also; his date of admission was May, 1891. This first List of Members contains the names and addresses of the 227 members of the Society.

Finkel and his co-editor Colaw, both members of the New York Mathematical Society, used its List of Members with much success in their campaign to obtain subscribers for their new journal. The first subscriber from a university was George Bruce Halsted at the University of Texas; he sent a check for \$30, and he contributed that amount annually as long as he was on the faculty at Texas. Finkel names [9, p. 310] the following as other university professors who became subscribers: E. H. Moore, Chicago; W. E. Byerly, Harvard; Edwin S. Crawley, Pennsylvania; Robert J. Aley, who forwarded sixteen subscriptions from Indiana; Irving Stringham, California; and H. A. Newton, Yale. Such university mathematicians as Halsted, L. E. Dickson, David Eugene Smith, and E. H. Moore contributed articles to early numbers and volumes of the *Monthly*.

Professor Finkel’s duties as a member of the faculty of Drury College were heavy [9, pp. 319–320] and the mathematical level of the *Monthly* rose because it was ignored by secondary school teachers but enthusiastically supported by college and university mathematicians. Finkel needed help, and he obtained it through a special relationship which developed with the University of Chicago and with E. H. Moore in particular, whom he called “my friend.” The nature and significance of this relation has been described by H. E. Slaught [17], who became in 1892 one of the first three fellows in mathematics at Chicago [18, p. 16], where he received his Ph.D. in 1898 and was a member of the faculty until 1931. In his appreciation of

Moore, Slaught wrote as follows [17, p. 193]:

But there is another phase of Professor Moore's remarkable influence in the general domain of mathematics which might easily be overlooked, but which should not be underestimated, amidst the greater effulgence of his brilliant research career; namely, his deep interest in the teaching of mathematics in the secondary schools and colleges and his ardent support of the Mathematical Association of America and of this MONTHLY as its official journal dedicated to the interests of mathematics in the collegiate field. It will be worthwhile to trace the development of this pedagogical interest in Moore's mind. It began in 1894 when he gave personal encouragement to Benjamin F. Finkel who was then just starting the AMERICAN MATHEMATICAL MONTHLY. They both hoped that this new journal would attract and inspire high school teachers. In this they were greatly disappointed but the new periodical soon found a fitting place in the collegiate world. Moore and Dickson contributed articles in the early numbers which gave encouragement to Finkel and prestige to the MONTHLY and later, in 1902, with Moore's advice and consent, Dickson accepted the co-editorship and the University of Chicago began an annual subsidy contribution of \$50.00 which was continued till the MONTHLY became self-supporting in 1915.

Moore was President of the American Mathematical Society in 1901 and 1902 [19, pp. 144–150], and his presidential retiring address [20], [17, pp. 193–194] emphasized his profound interest in the teaching of mathematics [21, pp. 169–171]. Moore's interest continued and took concrete form in action: in 1908 he urged Slaught to become co-editor of the *Monthly* when L. E. Dickson resigned to take up other editorial duties. Slaught thereafter became the visible leader in promoting the *Monthly* and the establishment of the MAA [22], [23], [24], but Moore's cooperation and counsel were of the greatest importance, especially in 1912 when the periodical was rescued from financial disaster and taken over by representatives of twelve universities and colleges in the middle west [25]. E. H. Moore was not present at the meeting in Columbus in December 1915 which organized the Mathematical Association of America, but there he was elected a member of its first Executive Council [1]. "His interest and satisfaction were approaching an upper bound when, in January 1916, the MONTHLY became the official journal of the newly organized Mathematical Association of America" [17, p. 194]. Slaught stated that Moore was "an habitual reader of the MONTHLY—problems and all. The first complete set of bound volumes of the periodical to be deposited in the Association library was the gift of Professor Moore and served as a token of his abiding interest in the movement for which the MONTHLY was to stand" [17, p. 194].

The Mathematical Association of America and *The American Mathematical Monthly* are an enduring monument to E. H. Moore's concern and support for the teaching of mathematics.

But the *Monthly* and the Mathematical Association of America derive their luster, not from their creation by the distinguished trio of Finkel, Slaught, and Moore, but rather from their devotion to the promotion of one of the great intellectual activities of civilization, namely, the study, teaching, and applications of mathematics.

Mathematics began in the dawn of history. There was mathematics in early Egypt and Mesopotamia; there was mathematics before the age of Greece and

Rome. Mathematics was an important part of the glory that was Greece, and educated people everywhere know the names of Pythagoras, Euclid, Archimedes, and Hypatia [34], [26]. The *Elements* of Euclid, written in Alexandria about 300 B.C., was the most influential textbook ever written. It has been translated into many languages. It was printed for the first time in Venice in 1482; estimates are that a thousand editions have been printed since then. Euclid might have written as follows about his *Elements*, using the words of the great Latin poet Horace [27]:

I have finished a monument more lasting than bronze and loftier than the pyramids reared by kings, that neither corroding rain nor the uncontrolled north wind can dash apart, nor the countless succession of years and the flight of ages. I shall not wholly die; that greater part of me shall escape Death and ever shall I grow, still fresh in the praise of posterity.

The beauty and the perfection of the structures and the relationships discovered by the mathematicians lead many to consider their subject the most recondite of the arts. G. H. Hardy wrote [28, p. 55], "I am interested in mathematics only as a creative art." Because of the beauty he saw in mathematics, Gauss called it the Queen of the Sciences [29, p. 1]. Edna St. Vincent Millay immortalized this beauty of mathematics in her sonnet entitled "Euclid Alone Has Looked on Beauty Bare" [30]; it reads as follows:

Euclid alone has looked on Beauty bare.
Let all that prate of Beauty hold their peace,
And lay them prone upon the earth, and cease
To ponder on themselves, the while they stare
At nothing, intricately drawn nowhere
In shapes of shifting lineage. Let geese
Gabble and hiss, but heroes seek release
From dusty bondage into luminous air.

Oh, blinding hour—oh, holy terrible day—
When first the shaft into his vision shone
Of light anatomized! Euclid alone
Has looked on Beauty bare; fortunate they
Who though once only, and then but far away,
Have heard her massive sandal set on stone.

Mathematics is not only an art form, but it is also an essential tool of the sciences [31]. For this reason, E. T. Bell called mathematics the Handmaiden of the Sciences [32]. Clarence R. Wylie, Jr. has emphasized this aspect of mathematics in his poem [33] entitled "Paradox":

Not truth, nor certainty. These I forswore
In my novitiate, as young men called
To holy orders must abjure the world.
"If . . . , then . . . ," this only I assert;
And my successes are but pretty chains
Linking twin doubts, for it is vain to ask
If what I postulate be justified,
Or what I prove possess the stamp of fact.

Yet bridges stand, and men no longer crawl
 In two dimensions. And such triumphs stem
 In no small measure from the power this game,
 Played with the thrice-attenuated shades
 Of things, has over their originals.
 How frail the wand, but how profound the spell!

Isaac Newton discovered the calculus in 1665–1666, published his *Philosophiae Naturalis Principia Mathematica* in 1687, and started the scientific developments which led to the Industrial Revolution in the eighteenth century. Riemann contributed his revolutionary treatment of geometry in 1854, which was used by Einstein in 1916 in his general theory of relativity, and many mathematicians contributed to the foundations which helped to establish quantum mechanics in this twentieth century.

The MAA serves a noble cause; the MAA participates in a noble tradition; the MAA is another of the special organizations and societies which, since ancient times, have been created to promote the study of mathematics. Pythagoras (about 580–500 B.C.), born on the island of Samos, established the communal and secret society of the Pythagoreans at Croton, on the southeastern coast of Italy. Many believe that this group developed most of the results in the first two books of Euclid's *Elements* [34, chap. IV]. Plato (about 428–348 B.C.) founded the Academy of Plato in Athens in 387 B.C.; considered by some to be the first university, it survived 900 years until A.D. 529. Plato is known, not as a mathematician, but as a maker of mathematicians; most of the important mathematicians of his time were associated with his Academy [34, chap. VI]. The inscription (in Greek) over the door of Plato's Academy, "Let no one unversed in geometry enter my doors," is preserved in the seal of the American Mathematical Society [19, p. ii]. The Museum of Alexandria, established by Ptolemy I shortly after 306 B.C., was another of the early institutions which promoted the study of mathematics; in particular, it sheltered Euclid, who flourished about 300 B.C. and was probably educated at Plato's Academy in Athens [34, chap. VII]. After the schools in Athens were closed in A.D. 529, many scholars moved east to the Arab world. Al-Mamun, Caliph of Baghdad (A.D. 809–833), established his "House of Wisdom," which was similar to the Museum at Alexandria. The important mathematician Mohammed ibn-Musa al-Khowarizmi was a member of the faculty of this "House of Wisdom" institution; his name, corrupted, and the title of one of his books have contributed the modern mathematical terms *algorithm* and *algebra* [34, pp. 251–258].

Copernicus and Galileo [35, chaps. II, III] brilliantly promoted the development of modern science in the sixteenth and seventeenth centuries, but the universities, long dominated by the Scholastics, gave them no support but actively opposed them [35, pp. 2, 35–38]. Thus there arose a need for new institutions to foster learning; the institutions created to fill this need were the scientific societies and the academies. The first formally organized society was the Accademia dei Lincei, founded in Rome in 1603 [36, p. 281]. Galileo became a member in 1611, but the academy faded out after Galileo's condemnation in 1633 (revived later, it became the national academy of Italy in 1875). Science had gained much momentum by the middle of the seventeenth century, and scientists supported and promoted their subject by organizing the following three new academies, which actively conducted research of many kinds: Accademia Del Cimento (Florence, 1657) [36, pp. 283–284]; Royal Society (London, 1662) [36, p. 359]; Académie Royale des Sciences (Paris, 1666) [36, pp. 304–305]. Although each academy included mathematics among its

interests, an organization devoted entirely to mathematics was to be expected; the first one organized which survives today was the Mathematische Gesellschaft, founded in Hamburg in 1690. This society is celebrating its tercentenary in 1990 [37], [38].

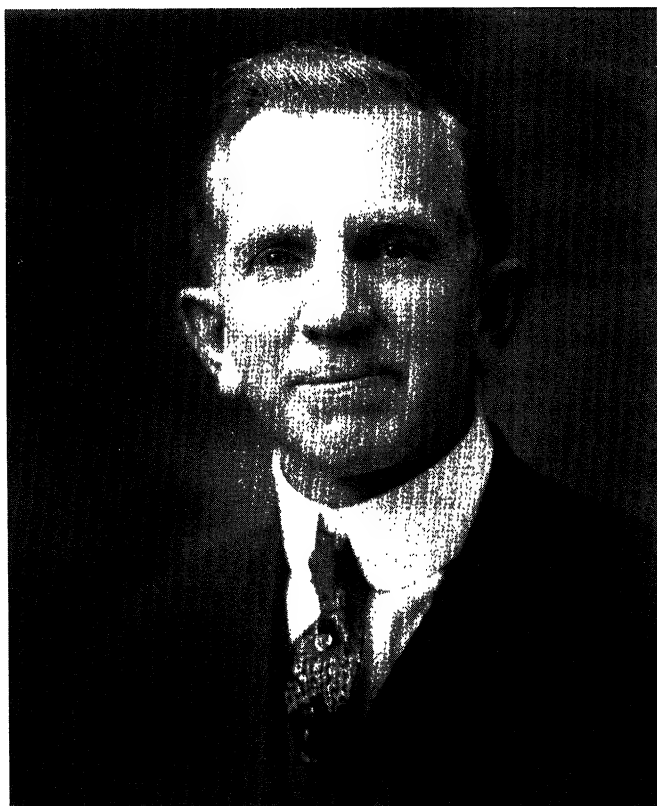
The societies and academies promoted their programs by publishing journals. The first periodical was the *Journal des Sçavans*, first published in Paris on January 5, 1665; it covered all learning, including science and mathematics [36, p. 305]. The first scientific periodical was *Philosophical Transactions*, first published by Oldenburg, the Secretary of the Royal Society, on March 6, 1665, and taken over by the Society in 1753 [36, p. 359]. After these beginnings, societies, academies, and their publications increased slowly during the eighteenth century, but more rapidly during the nineteenth century. Bartle [37] states that “the first journal devoted exclusively to mathematics was the French journal *Annales de Mathématique Pures et Appliquées*, founded and edited by J. D. Gergonne and published from 1810 to 1831.” Bartle also describes other early mathematics journals, and he summarizes their growth in number as follows:

According to Müller, before 1700 there were only 17 journals publishing articles with mathematical content, during the eighteenth century there were 210 new journals with mathematical articles, and during the nineteenth century there were another 950 new journals. (In this reckoning a continuation of a journal under a different title is considered a ‘new’ journal.) However, before 1900 about half of these journals had ceased publication or changed titles, so that in 1900 there were about 600 existing journals containing some mathematical articles!

R. C. Archibald’s article [38] entitled “Mathematical Societies and Periodicals” in the fourteenth edition of the Encyclopaedia Britannica contains a brief history of these developments. Archibald states that “the great national mathematical societies were established in their countries in the following order: Russia, Great Britain, France, Italy, United States, and Germany.” Most of the mathematical journals publish research articles, but a few are devoted to the history of mathematics and to the teaching of mathematics. For example, Archibald states that in Great Britain “the Mathematical Association (about 1,160 members) was founded in 1871 as the Association for the Improvement of Geometrical Teaching, and took its present name in 1897. It has published *Reports* (1871–93) and *Mathematical Gazette* (1894 +).” Archibald, a member of the MAA’s first Executive Council (see above), reported as follows on the MAA and its journal in 1929: “The Mathematical Association of America, founded in 1915, and now having over 2,000 members, aspires particularly to serve the colleges of the country by awakening and sustaining interest in mathematics and by fostering the beginnings of mathematical research. Its official organ (1916 +) is *The American Mathematical Monthly* (1894 +), founded and published for many years by B. F. Finkel.”

Thus special societies and institutions have supported mathematics throughout history, and the MAA is honored today to be included as one of this illustrious company.

Civilization began with the Stone Age, and it progressed through the Bronze Age, the Iron Age, and the Steel Age. In 1915 civilization arrived at the beginning of the Age of the Mathematical Association of America (1915–1990). But the growth of mathematics and its applications has converted the Age of the MAA into the Age of the Mathematical Sciences. In 1915 the American Mathematical



B. F. Finkel
Editor of *The American Mathematical Monthly* 1894–1912

Society was the only professional organization in the field of mathematics in America. Since that time it has been joined not only by the Mathematical Association of America but also by the Canadian Mathematical Society, the National Council of Teachers of Mathematics, the Association for Symbolic Logic, the Institute of Mathematical Statistics, the Society for Industrial and Applied Mathematics, the Association for Computing Machinery, and the Operations Research Society of America. The establishment of these new professional organizations emphasizes the growth of mathematics and the increasing number of areas of specialization and their applications. In addition to the new professional organizations, America now has many strong departments of mathematics in universities and in government and industrial laboratories.

Four great ages of civilization—Stone, Bronze, Iron, and Steel—preceded the founding of the Mathematical Association of America, and in 1915 the Radio Age was just beginning. The age of the MAA has brought a succession of spectacular ages which make it the most fantastic age in all of history. The beginning of the Nuclear Age burst upon the world in 1945. The Computer Age began with computers built during World War II. Television was developed shortly before World War II, and the Television Age began with widespread commercial broadcasting around 1950. The Space Age began on October 4, 1957, when the Russians put Sputnik I in orbit, and it reached a certain maturity when men landed on the moon July 1969. The Jet Age began in 1958 when the Boeing 707 jet airplane

began regular commercial passenger service. Finally, the Age of the MAA includes the beginning of the Age of Molecular Biology.

The Mathematical Association of America has actively participated in these great events in the Age of the Mathematical Sciences. Mathematics is no longer an elective for civilization; it is required for the proper functioning of society. The MAA maintains a program designed to satisfy these needs. It holds two national meetings each year with the American Mathematical Society and other organizations; these meetings present lectures, committee reports and panel discussions, and mini-courses of instruction in new subjects and applications. The MAA publishes three journals and a newsletter; its book catalog lists the titles of more than a hundred books in print. The MAA maintains a program of visiting lecturers; it conducts talent searches through contests at several levels; and major efforts have gone into curriculum studies which covered the mathematical sciences broadly [39]. The MAA has special programs designed to promote the study of the mathematical sciences by women and minorities. Finally, in this its seventy-fifth year, the Mathematical Association of America finds itself back where Finkel began in 1894: it is deeply involved—with many others—in efforts to improve the teaching of mathematics at all levels. The improvement of teaching is a continuing problem, but the need is especially great today.

And so let us celebrate the seventy-fifth anniversary of the founding of the Mathematical Association of America. The group of 104 men and women who met here in 1915 planned well and wisely; the organization which they established has developed into a major learned society of almost 31,000 members. Let us celebrate with lectures and with publications. Let us celebrate by renewing our efforts to share the Beauty and the Power of mathematics with all people.

REFERENCES AND NOTES

1. W. D. Cairns, The Mathematical Association of America, *The American Mathematical Monthly*, 23 (1916), 1–6.
2. T. H. Gladstone, *The Englishman in Kansas*, Miller & Company, New York, 1857. Republished by the University of Nebraska Press, Lincoln, Nebraska, 1971. lxvi + 328 pages. Thomas H. Gladstone, a kinsman of the distinguished English statesman, William Ewart Gladstone, was a journalist and a correspondent for *The Times* of London. He came to the United States during the latter part of the winter of 1855–56, and he was in Kansas from May 22, 1856—the day after the pro-slavery forces had made a devastating assault on Lawrence—until early 1857. This book is based on his stay in the United States, on his visit to Kansas, and on his dispatches to *The Times*.
3. Leverett Wilson Spring, *Kansas: The Prelude to the War for the Union*, Houghton Mifflin Company, Cambridge, revised edition, 1906, viii + 340 pages. The first edition of this book was written in 1885 while Professor Spring was a member of the faculty of The University of Kansas.
4. Carl W. Breihan, *The Complete and Authentic Life of Jesse James*, F. Fell, New York, 1954, 287 pages.
5. *History of Clinton and Caldwell Counties, Missouri*. Clinton County by Carrie Polk Johnston, and Caldwell County by W. H. S. McGlumphy, Historical Publishing Company, Topeka and Indianapolis, 1923, 836 pages.
6. *History of Caldwell and Livingston Counties, Missouri*, National Historical Company, St. Louis, 1886.
7. Kidder, a small town in the violent and lawless Missouri-Kansas border (see [2], [3], [4], [5], [6]) seems an unlikely environment to support an important intellectual enterprise; yet Finkel—a mathematician with a mission—defied the odds and successfully launched *The American Mathematical Monthly* there in 1894.
8. Biography of Benjamin Franklin Finkel, *American Men of Science*, sixth edition, New York, The Science Press, 1938.
9. B. F. Finkel, The human aspect in the early history of the American Mathematical Monthly, *American Mathematical Monthly*, 38 (1931), 305–320.

10. B. F. Finkel, *A Mathematical Solution Book, Containing Systematic Solutions of Many of the Most Difficult Problems, with Notes and Explanations*, fourth edition, revised and enlarged, Kibler & Company, Publishers, Springfield, Missouri, 1902. xvi + 549 pages. The first edition of this book was published in Kidder, Missouri in 1893.
11. *History of Greene County, Missouri*, Western Historical Company, St. Louis, 1883.
12. Harold Bell Wright, *The Shepherd of the Hills*, Grosset & Dunlap, xi + 299 pages. First copyright 1907.
13. Garrett Birkhoff, *Mathematics at Harvard, 1836–1944. A Century of Mathematics in America*, Part II, pages 3–58. American Mathematical Society, History of Mathematics, vol. 2, 1989, x + 585 pages.
14. Dirk J. Struik, The MIT Department of Mathematics During Its First Seventy-Five Years: Some Recollections. *A Century of Mathematics in America*, Part III, pages 163–177. American Mathematical Society, History of Mathematics, vol. 3, 1989, ix + 675 pages.
15. H. W. Tyler, Biography of John Daniel Runkle, *American Mathematical Monthly*, 10 (1903), 183–185.
16. Simon Newcomb, An account of Professor Runkle's *Mathematical Monthly*, *American Mathematical Monthly*, 10 (1903), 130–133.
17. H. E. Slaught, Eliakim Hastings Moore: An Appreciation by H. E. Slaught, *American Mathematical Monthly*, 40 (1933), 191–195.
18. Kenneth O. May, editor, *The Mathematical Association of America: Its First Fifty Years*, Mathematical Association of America, Washington, D.C., 1972, vii + 172 pages.
19. Raymond Clare Archibald, *A Semicentennial History of the American Mathematical Society*, New York, American Mathematical Society, 1938, xiv + 262 pages.
20. E. H. Moore, On the foundations of mathematics, *Bulletin of the American Mathematical Society*, 9 (1903), 402–424.
21. Karen Hunger Parshall, Eliakim Hastings Moore and the Founding of a Mathematical Community in America, 1892–1902. *A Century of Mathematics in America*, Part II, pages 155–175. American Mathematical Society, History of Mathematics, vol. 2, 1989, x + 585 pages.
22. H. E. Slaught, Retrospect and prospect, *American Mathematical Monthly*, 19 (1912), 183–186.
23. H. E. Slaught, The promotion of collegiate mathematics, *American Mathematical Monthly*, 22 (1915), 251–253.
24. H. E. Slaught, The teaching of mathematics, *American Mathematical Monthly*, 22 (1915), 289–292.
25. B. F. Finkel, Reminiscences and appreciations, in Notes and News, *American Mathematical Monthly*, 19 (1912), 197–201.
26. Margaret Alic, *Hypatia's Heritage: A History of Women in Science from Antiquity through the Nineteenth Century*, Beacon Press, Boston, 1986, x + 230 pages.
27. Horace, Odes, 3, xxx. The original in Latin can be found on page 129 of the following book: Charles E. Bennett, *Horace: Odes and Epodes*, Allyn and Bacon, Boston, New York, Chicago, 1901, xl + 406 pages. The translation quoted in the paper above appears on page 307 of the following book: E. T. Bell, *Men of Mathematics*, Simon and Schuster, New York, 1937, xxi + 592 pages.
28. G. H. Hardy, *A Mathematician's Apology*, Cambridge, At the University Press, 1940. vii + 93 pages.
29. E. T. Bell, *The Queen of the Sciences*, The Williams & Wilkins Company, Baltimore, 1931. iii + 138 pages.
30. Edna St. Vincent Millay, "Euclid Alone Has Looked on Beauty Bare." From *The Ballad of the Harp-Weaver and Other Poems*, Harper and Brothers, 1920. Also found on page 1162 of the following book: James Dow McCallum, editor, *The 1936 College Omnibus*, Harcourt, Brace and Company, New York, 1936, x + 1193 pages.
31. Eugene P. Wigner, The Unreasonable Effectiveness of Mathematics in the Natural Sciences, pages 123–140 of the following book edited by Thomas L. Saaty and F. Joachim Weyl: *The Spirit and the Uses of the Mathematical Sciences*, McGraw-Hill Book Company, New York, 1969, x + 301 pages.
32. Eric Temple Bell, *The Handmaiden of the Sciences*, The Williams and Wilkins Company, Baltimore, 1937, viii + 216 pages.
33. The poem entitled "Paradox," by Clarence R. Wylie, Jr., appeared in the July, 1948 issue of *Scientific Monthly*. It is printed also on page iii of the following book: Richard B. Kershner and L. R. Wilcox, *The Anatomy of Mathematics*, The Ronald Press Company, New York, 1950, xi + 416 pages.
34. Carl B. Boyer, *A History of Mathematics*, John Wiley & Sons, Inc., New York, London, Sydney, 1968, xv + 717 pages.

35. A. Wolf, *A History of Science, Technology and Philosophy in the 16th and 17th Centuries*, The Macmillan Company, New York, 1935, xxvii + 692 pages. With 316 illustrations.
36. Robert Mortimer Gascoigne, *A Chronology of the History of Science, 1450–1900*, Garland Publishing, Inc., New York & London, 1987, xi + 585 pages.
37. Robert G. Bartle, A Brief History of the Mathematical Literature, pages 3–6 in *Mathematical Reviews Special Issue, 50th Anniversary Celebration*, 18 January 1990, 23 pages.
38. Raymond Clare Archibald, Mathematical Societies and Periodicals, *Encyclopedia Britannica*, fourteenth edition, 1929, volume 15, pages 75–78.
39. *A Compendium of CUPM Recommendations*, Published by the Mathematical Association of America, Washington, D.C. The individual reports carry the dates on which they were issued. Volume I, iv + pages 1–457; Volume II, iv + pages 458–756. Each volume has a supplement which lists all of those who were members of CUPM panels and subcommittees.

Department of Mathematics
The University of Kansas
Lawrence, Kansas 66045-2142

The MONTHLY closes its third year since its reorganization with many of its earlier fears and anxieties removed, with its subscription list more than trebled, with a large and increasing body of contributors, and with a feeling of certainty that no mistake was made in regard to the opportunity that seemed to be offered in the field in which its operations have been centered. High praise is due to those colleges and universities which deemed the cause represented by the MONTHLY to be worthy of their consideration and which made possible the promotion of this cause by their generous subsidy contributions throughout the three-year period. Hearty thanks are due the representatives of these institutions who have constituted the Board of Editors and who have unselfishly given their time and their services, amounting in the aggregate to no small consideration, thus showing their faith in this cause and their willingness to back it by works. Full credit is due the constituency of the MONTHLY, a widely representative body of persons who have shown their confidence and indicated their loyalty by their continued adherence in ever-increasing numbers. Finally, and not least, credit is due those who have made their contributions to the cause by purchasing advertising space, and to The New Era Printing Company, whose excellent composition and presswork are worthy of mention.

—*American Mathematical Monthly*
 22, (1915) p. 352.

An Axiomatic Approach to the Integral

Leonard Gillman

1. THE RIEMANN INTEGRAL. Have you ever watched an engineer or physicist set up an integral? They don't mention Riemann sums nor pick an arbitrary point z_k in the k th segment; they don't even mention a partition. Instead, they draw FIGURE 1 (to use area as an example) and say: *Here's the strip at x of width dx (where dx is small). Then dA is equal to* [writing it down]

$$f(x) dx.$$

Then they prefix an integral sign.

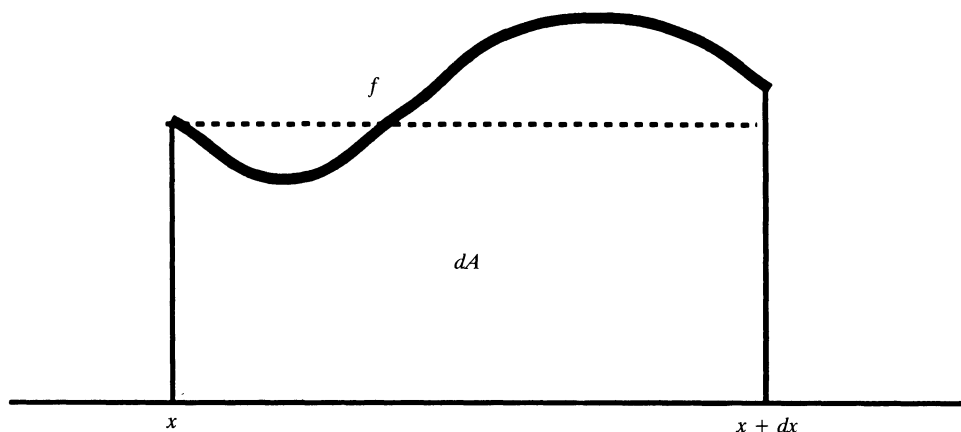


Figure 1

This invariably gets a good laugh at a lecture. Maybe instead it deserves applause for being so efficient and logical. Concentrating on a single segment permits circumventing the notational paraphernalia of all those subscripts. Calling the width dx permits skipping the summation sign and going directly to the integral sign. (Some concerned teachers will argue that it is better to remind the student of the sums on which the integral is based. Others will liken that to solving quadratics by completing the square.)

The usual way is to pick an arbitrary z_k in the k th segment of a partition, form the corresponding Riemann sum, and proclaim that since arbitrary Riemann sums approach the integral (as the norm goes to 0), so then do these arbitrary sums. But then so do particular sums. The practicing scientist picks $z_k = x_{k-1}$. (It is assumed in all this that f is continuous.)

Actually, there are applications in which one traditionally chooses a particular z_k —for example, in length of arc, where z_k is determined from the Mean-Value Theorem. (By that time the students are wondering whether it's legal.) Our scientist sticks to $z_k = x_{k-1}$ here as well. Presumably the resulting partial sums also approach “what should be” or “what the physicists intuition affirms is” the volume, or the work, etc.

2. THE DARBOUX INTEGRAL. Can this hope be replaced by a theorem? We suggest an approach based on the Darboux integral, defined as the unique number lying between all lower sums and all upper sums. Probably every calculus text bases at least one of its derivations on this condition. We will exploit it systematically. Our point of departure is the well-known properties of additivity and “betweenness”:

$$(A) \quad \int_a^{x+\Delta x} \varphi = \int_a^x \varphi + \int_x^{x+\Delta x} \varphi,$$

$$(B) \quad \left(\min_{[x, x+\Delta x]} \varphi \right) \Delta x \leq \int_x^{x+\Delta x} \varphi \leq \left(\max_{[x, x+\Delta x]} \varphi \right) \Delta x$$

(where φ is continuous on an interval $[a, b]$, and $a \leq x < x + \Delta x \leq b$).

It should come as no surprise that the integral is the *only* function satisfying (A) and (B); this is stated formally as Theorem 1 below. In each application, to show that the quantity of interest is an integral, we show that it has these two properties. As a result, our intuition is relieved of the responsibility of making predictions about infinite processes. (We don't even mention Riemann sums.) Instead our assumptions refer to concepts that are more real to the student, such as that the area of an enclosing region is greater than that of the enclosed region; and we put these assumptions up front. Integrals are thus derived as *theorems* rather than announced as definitions. As a byproduct, we know that the volume by either discs or shells will be the same, that the area by either rectangular or polar coordinates will be the same, and so on.

In the usual treatment of the integral, after all the elaborate preparation, there always comes that anticlimactic moment when you confront a nonRiemann sum and have to mumble that yes, that will work too but the proof is too hard for this course. (The rough-and-ready scientist skirts around this problem.) Our more general Theorem 2 covers these cases as well.

Most of all this has been done before in various degrees of thoroughness, but a review seems worth while. What I believe to be new are the improved version of the general theorem, a more natural axiom for surface area, and the observation that this axiom and those for arc length are equivalent (in the presence of additivity) to the corresponding formulas.

3. BACKGROUND

Theorem 1. Let φ be continuous on an interval $[a, b]$, and let I_u^ν be defined for $a \leq u \leq \nu \leq b$. Suppose that

$$(A) \quad I_a^{x+\Delta x} = I_a^x + I_x^{x+\Delta x}$$

and

$$(B) \quad \left(\min_{[x, x+\Delta x]} \varphi \right) \Delta x \leq I_x^{x+\Delta x} \leq \left(\max_{[x, x+\Delta x]} \varphi \right) \Delta x,$$

when $\Delta x > 0$. Then

$$I_a^b = \int_a^b \varphi(x) dx.$$

Note that (A) includes the condition $I_a^a = 0$ (the case $x = a$).

Proof: Consider an arbitrary partition of $[a, b]$. By hypothesis, (B) holds for every segment $[x, x + \Delta x]$. Hence by (A) (applied repeatedly), the lower and upper sums L and U satisfy $L \leq I_a^b \leq U$. Therefore I_a^b is the Darboux integral.

The earliest reference I know to the properties (A) and (B) as characteristic of the integral is Hahn and Rosenthal [2, 149–150], which in fact adopts them as the *definition*. I believe that Howard Levi [4, 60–70] is the first to recognize that they characterize the integral and then follow up by invoking them systematically in applications; this book, with its wealth of imaginative ideas, deserves to be better known. Serge Lang [3] also invokes this characterization. Gillman and McDowell [1] adopts (A) and (B) as the definition of the integral and applies them systematically to an extensive selection, including polar coordinates and multiple integrals. The two-variable version of Theorem 1 permits an effortless proof that the value of the double integral is given by each of the two iterated integrals.

This illustrates the advantage in using (A) and (B) rather than the Darboux integral itself, which is essentially the same thing: they focus attention on the underlying principles. (Recall that (A) and (B) constitute the two steps of the proof of the Fundamental Theorem of Calculus.)

We now illustrate some standard geometric applications. The letters A, V, L, S , with appropriate indices, represent area, volume, arc length, surface area.

4. APPLYING THEOREM 1. The bread-and-butter problems are (a) the area under the graph of a function f , and (b) the volume generated by revolving the graph of f about the x -axis. Most textbooks use a betweenness argument in these two problems, so for them we describe our method only in outline. In (a), we bound the area on $[x, x + \Delta x]$ by two rectangular regions on that same base, of heights $\min f$ and $\max f$; then we use the fact that the area of a rectangle is the product of its dimensions, and the axiom about the areas of enclosed and enclosing regions. In (b), we bound the volume on $[x, x + \Delta x]$ by two right circular cylinders, of radii $\min f$ and $\max f$; then we use the fact that the volume of a right circular cylinder is $\pi r^2 h$, and the axiom that if one of two solids encloses another, the enclosing one has the larger volume.

We proceed to examples where the bounds are less obvious.

Example 1. Length of arc. Let f and f' be continuous on $[a, b]$. To define the length of the graph of f , first note axiom (A): $L_a^{x+\Delta x} = L_a^x + L_x^{x+\Delta x}$.

Now we look at FIGURE 2 and wonder what to do. The chord of the arc is a lower bound for the length; what is an upper bound? A cue is the observation that of two straight-line graphs on the same interval, the one with the larger absolute slope is the longer (FIGURE 3). We extend this principle to graphs with variable slopes. Imagine walking across a field, heading eastward but edging north as you go. Suppose I do the same, keeping due north of you at all times and, *at every instant*, heading *more* northward than you (FIGURE 4). Everyone agrees I walk farther than you. Also, south is as good as north. We summarize this in the axiom:

(L) If $|f'_1(t)| \leq |f'_2(t)|$ for each t in $[x, x + \Delta x]$, then $L_x^{x+\Delta x}(f_1) \leq L_x^{x+\Delta x}(f_2)$.

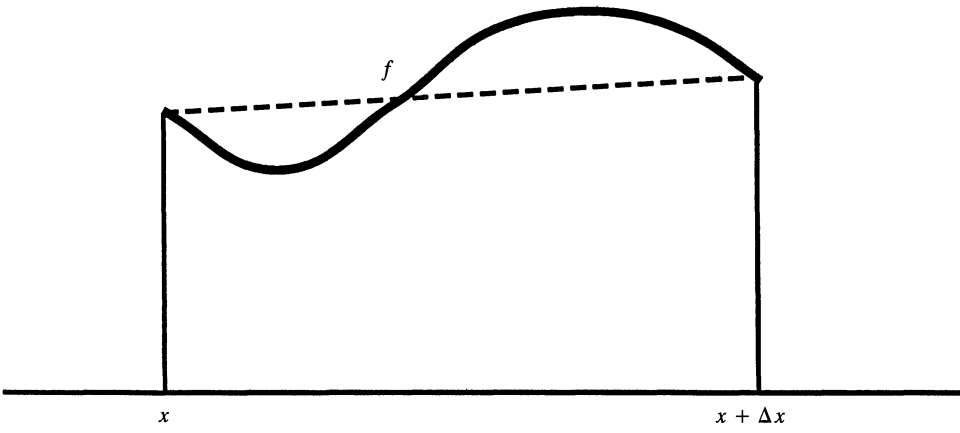


Figure 2



Figure 3

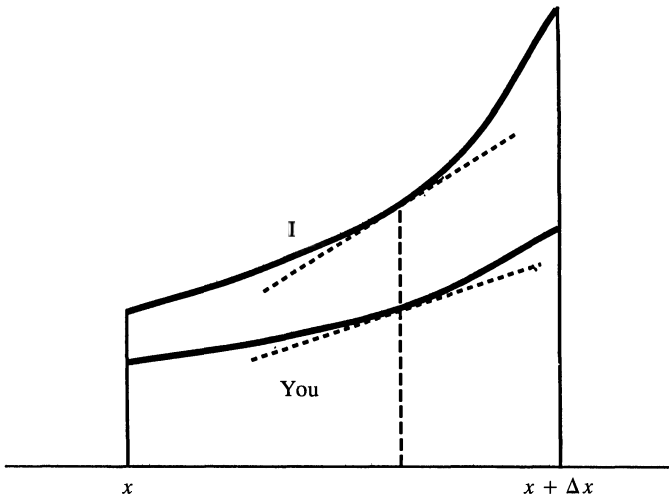


Figure 4

According to this axiom, the segment on $[x, x + \Delta x]$ with slope $\min|f'|$ is shorter than the graph, which in turn is shorter than the segment with slope $\max|f'|$. (Here and later we could state the axiom for the special case and then it would look less forbidding.) Noting that the length of a segment of slope m is $\sqrt{1 + m^2} \Delta x$, and that $\sqrt{1 + \min|f'|^2} = \min\sqrt{1 + f'^2}$ and similarly for the max, we have

$$\min_{[x, x + \Delta x]} \sqrt{1 + f'^2} \Delta x \leq L_x^{x + \Delta x} \leq \max_{[x, x + \Delta x]} \sqrt{1 + f'^2} \Delta x. \quad (1)$$

This is property (B) with respect to the function $\varphi(x) = \sqrt{1 + f'(x)^2}$. By Theorem 1,

$$L_a^b = \int_a^b \sqrt{1 + f'(x)^2} dx. \quad (2)$$

Note that we never did use the original chordal lower bound. The weaker lower bound in (1) is adequate, and its value is expressible directly, without the Mean-Value Theorem.

Axiom (L) is stated and applied in [4] and again in [1]. It may be that the notion of approximating an arc by chords holds such a strong intuitive appeal that the foregoing derivation may be dispensed with. In any case, we will want these ideas in the discussion of surface area, where intuition tends to be weak.

5. SETTING UP INTEGRALS. Let us recapitulate. To set up an integral on $[a, b]$, consider a typical subinterval $[x, x + \Delta x]$. Denote by I_a^b the quantity in the application to be represented by the integral. First verify the properties (A) and (B) *within the field of application*—on the basis of knowledge of the field, or intuition, or advice from a physicist, etc. *From then on, the rest is mathematics*: Theorem 1 tell us that I_a^b is the integral.

To verify (A) and (B): (i) Draw a picture. (ii) Verify (A). This is an *axiom* in the application and turns out to be automatic in just about every case. (iii) Find a formula that holds when all the variables are constants. (iv) Use this formula, along with additional axioms for the application, to obtain the bounds to be used in (B). Note that (B) is a *theorem* in the application.

6. THE GENERAL THEOREM. Theorem 1 proves to be inadequate in many applications, as we will shortly see. We need the following generalization.

Theorem 2. *Let φ be continuous on $[a, b]$ and let I_u^v be defined for $a \leq u \leq v \leq b$. Suppose that for $\Delta x > 0$,*

$$(A) \quad I_a^{x + \Delta x} = I_a^x + I_x^{x + \Delta x}$$

and

$$(B') \quad \alpha \Delta x \leq I_x^{x + \Delta x} \leq \beta \Delta x,$$

where α and β both approach $\varphi(x)$ as $\Delta x \rightarrow 0$. Then

$$I_a^b = \int_a^b \varphi(x) dx.$$

Proof: We show as in the Fundamental Theorem that the function

$$F(x) \equiv I_a^x$$

is an antiderivative of φ . (The notation will be for the case $\Delta x > 0$.) By (A),

$$\frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{I_x^{x + \Delta x}}{\Delta x}.$$

By (B),

$$\alpha \leq \frac{I_x^{x+\Delta x}}{\Delta x} \leq \beta.$$

Since α and β approach $\varphi(x)$ as $\Delta x \rightarrow 0$, so does the quantity between them; therefore $F'(x) = \varphi(x)$. Finally, by the other half of the Fundamental Theorem,

$$\int_a^b \varphi(x) dx = I_a^b - I_a^a = I_a^b.$$

A special case of this theorem appears in [4]. The full theorem is presented in [1] for both the single and double integral, but the statements are somewhat awkward and the proof for the double integral is ε 's and δ 's; I have since written up a proof for myself modeled after the one just given.

7. APPLYING THE GENERAL THEOREM. The procedure for setting up integrals is the same as that outlined in Section 5, except possibly in step (iv), where there may be more than one choice for α and β .

Example 2. Area between two graphs. Let f and g be continuous on $[a, b]$, with $f \geq g$. To define the area between their graphs, I would first add the constant $|\min_{[a,b]} g|$ to both functions (if necessary) to reduce to the case $g(x) \geq 0$ for all x (with the axiom that the area between the graphs is not affected by the rigid motion), then note (by an axiom of general additivity) that the area between f and g is the area under f minus the area under g . But the very simplicity of the problem, free of distractions, makes it a good one for illustrating our methods. Again, we first note axiom (A): $A_a^{x+\Delta x} = A_a^x + A_x^{x+\Delta x}$.

Consider the strip on $[x, x + \Delta x]$ (FIGURE 5). Is it an axiom that the area between the graphs is greater than that of a rectangle on the same base with height $\min_{[x, x+\Delta x]}(f - g)$, and less than one of height $\max_{[x, x+\Delta x]}(f - g)$? Note that in this example, the region does not *contain* any rectangle on that base of height

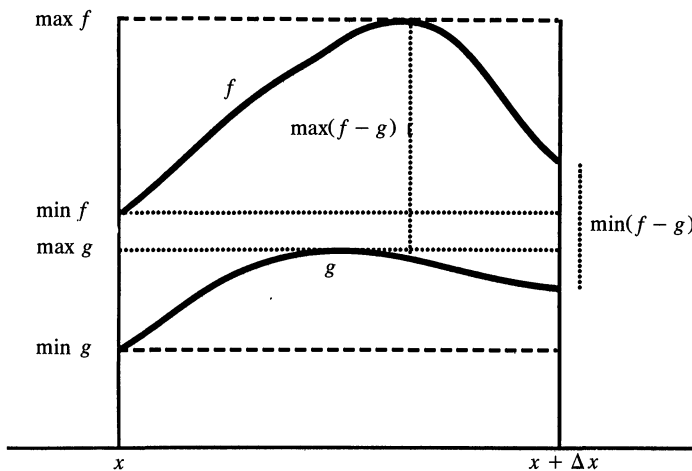


Figure 5

$\min_{[x, x+\Delta x]}(f - g)$, nor is it contained in any rectangle of height $\max_{[x, x+\Delta x]}(f - g)$. If you want that axiom nevertheless, then you obtain

$$\min_{[x, x+\Delta x]} (f - g) \Delta x \leq A_x^{x+\Delta x} \leq \max_{[x, x+\Delta x]} (f - g) \Delta x. \quad (3)$$

This is property (B), with $\varphi = f - g$, and you are finished.

However, the general theorem does not require your intuition to be that fine. You need only compare regions of which one is a subset of the other. To get an upper bound in the present example, pick the enclosing rectangle of height

$$\max_{[x, x+\Delta x]} f - \min_{[x, x+\Delta x]} g.$$

The lower bound comes with a little twist. If $\min_{[x, x+\Delta x]} f > \max_{[x, x+\Delta x]} g$, then the strip encloses a rectangle of height

$$\min_{[x, x+\Delta x]} f - \max_{[x, x+\Delta x]} g$$

(FIGURE 5); otherwise, $\min_{[x, x+\Delta x]} f - \max_{[x, x+\Delta x]} g$ is zero or negative. In either case,

$$\left(\min_{[x, x+\Delta x]} f - \max_{[x, x+\Delta x]} g \right) \Delta x \leq A_x^{x+\Delta x} \leq \left(\max_{[x, x+\Delta x]} f - \min_{[x, x+\Delta x]} g \right) \Delta x.$$

These inequalities are weaker than (3) and are not in the form (B). However, each of the two bounds (in parentheses) approaches $f(x) - g(x)$ as $\Delta x \rightarrow 0$. The inequalities are therefore of the form (B'), with $\varphi = f - g$. By the general theorem,

$$A_a^b = \int_a^b (f - g).$$

Example 3. Volume of a solid of revolution; shell method. Let f and g be continuous on $[a, b]$, where $a \geq 0$, and $f(x) \geq g(x)$ for all x , and revolve the region between the graphs about the y -axis to generate a solid. To define its volume by the shell method, we again first note axiom (A): $V_a^{x+\Delta x} = V_a^x + V_x^{x+\Delta x}$.

Consider the strip on $[x, x + \Delta x]$ (FIGURE 5). Its contribution to the total volume is less than that from the enclosing rectangle of height $\max_{[x, x+\Delta x]} f - \min_{[x, x+\Delta x]} g$, and is greater than that from the enclosed rectangle of height $\min_{[x, x+\Delta x]} f - \max_{[x, x+\Delta x]} g$ (perhaps a "negative" rectangle). The solid created by rotating a rectangle is the difference between two cylinders; the formula for the volume is $V = 2\pi \bar{r} h \Delta r$, where Δr is the difference of the two radii, \bar{r} is their average, and h is the height of the rectangle. Again we invoke the axiom that the larger volume goes with the enclosing solid. Consequently,

$$2\pi \bar{x} \left(\min_{[x, x+\Delta x]} f - \max_{[x, x+\Delta x]} g \right) \Delta x \leq V_x^{x+\Delta x} \leq 2\pi \bar{x} \left(\max_{[x, x+\Delta x]} f - \min_{[x, x+\Delta x]} g \right) \Delta x,$$

where $\bar{x} = x + \frac{1}{2} \Delta x$, the average radius. These inequalities are in the form (B'), with $\varphi(x) = 2\pi x[f(x) - g(x)]$. By the general theorem,

$$V_a^b = 2\pi \int_a^b x[f(x) - g(x)] dx.$$

Example 4. Length of an arc defined parametrically. For an arc defined by parametric equations $x = f(t)$, $y = g(t)$, where f , f' , g , and g' are continuous on an interval $a \leq t \leq b$, the idea is the same as in Example 1. Again we note axiom (A): $L_a^{t+\Delta t} = L_a^t + L_t^{t+\Delta t}$.

Consider the strip on $[t, t + \Delta t]$. Will you accept as an axiom that:

$$(LL_0) \quad \text{If } f_1'(\tau)^2 + g_1'(\tau)^2 \leq f_2'(\tau)^2 + g_2'(\tau)^2 \text{ for each } \tau \text{ in } [t, t + \Delta t], \\ \text{then } L_t^{t+\Delta t}(f_1, g_1) \leq L_t^{t+\Delta t}(f_2, g_2)?$$

If so, then you obtain

$$\min_{[t, t+\Delta t]} \sqrt{f'^2 + g'^2} \Delta t \leq L_t^{t+\Delta t} \leq \max_{[t, t+\Delta t]} \sqrt{f'^2 + g'^2} \Delta t.$$

This is property (B), with $\varphi(t) = \sqrt{f'(t)^2 + g'(t)^2}$, and you are finished.

The general theorem does not require you to be so imaginative. Instead, consider the following more pedestrian axiom:

$$(LL) \quad \text{If } |f_1'(\tau)| \leq |f_2'(\tau)| \text{ and } |g_1'(\tau)| \leq |g_2'(\tau)| \text{ for each } \tau \text{ in } [t, t + \Delta t], \\ \text{then } L_t^{t+\Delta t}(f_1, g_1) \leq L_t^{t+\Delta t}(f_2, g_2).$$

This leads to the inequalities

$$\sqrt{\min_{[t, t+\Delta t]} f'^2 + \min_{[t, t+\Delta t]} g'^2} \Delta t \leq L_t^{t+\Delta t} \leq \sqrt{\max_{[t, t+\Delta t]} f'^2 + \max_{[t, t+\Delta t]} g'^2} \Delta t,$$

which is property (B'), with $\varphi(t) = \sqrt{f'(t)^2 + g'(t)^2}$. By the general theorem,

$$L_a^b = \int_a^b \sqrt{f'(t)^2 + g'(t)^2} dt. \quad (4)$$

Note that this argument carries over at once to the three-dimensional case $x = f(t)$, $y = g(t)$, $z = h(t)$.

Example 5. Area of a surface of revolution. Let f and f' be continuous on $[a, b]$, with f nonnegative. When the graph of f is revolved about the x -axis, it generates a surface. To define the area of such a surface, we first note axiom (A): $S_a^{x+\Delta x} = S_a^x + S_x^{x+\Delta x}$.

Now we need information about some basic surface of revolution. The simplest is the one obtained by revolving a horizontal segment. If the length of the segment is h and the radius of revolution is r , then the segment sweeps out a right circular cylinder of radius r and "height" h ; its area is $2\pi rh$ (found by the "slit-and-unroll" procedure).

Consider two such cylinders with parameters r_1, h_1 and r_2, h_2 ; obviously, if $r_1 h_1 > r_2 h_2$, then the area of the first is greater than the area of the second. We wish to find a similar comparison for revolving any two graphs. Perhaps you feel confident that:

$$(S_0) \quad \text{If } f_1(t) \sqrt{1 + f_1'(t)^2} \leq f_2(t) \sqrt{1 + f_2'(t)^2} \text{ for each } t \text{ in } [x, x + \Delta x], \\ \text{then } S_x^{x+\Delta x}(f_1) \leq S_x^{x+\Delta x}(f_2).$$

If so then you obtain

$$2\pi \min_{[x, x+\Delta x]} (f \cdot \sqrt{1 + f'^2}) \Delta x \leq S_x^{x+\Delta x} \leq 2\pi \max_{[x, x+\Delta x]} (f \cdot \sqrt{1 + f'^2}) \Delta x.$$

This is property (B), with $\varphi(t) = f(t) \sqrt{1 + f'(t)^2}$, and you are finished.

But again your intuition does not have to be that creative. Instead, consider the cruder assumption that if the graph of f_1 lies below the graph of f_2 , and if

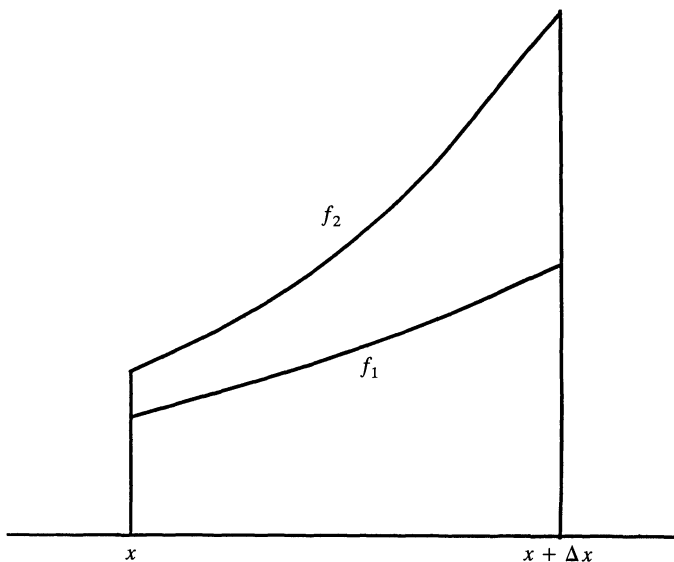


Figure 6

$|f'_1| \leq |f'_2|$ pointwise (FIGURE 6), then f_1 generates the smaller area:

- (S) *If $f_1(t) \leq f_2(t)$ and $|f'_1(t)| \leq |f'_2(t)|$ for each t in $[x, x + \Delta x]$,
then $S^{x+\Delta x}_x(f_1) \leq S^{x+\Delta x}_x(f_2)$.*

This axiom appears in [1] in a slightly different form. The present version follows along the lines suggested by the referee as being more intuitive.

We return to the given function f . Its graph is (say) as in FIGURE 7. The two accompanying line segments serve as lower and upper bounds: the lower has

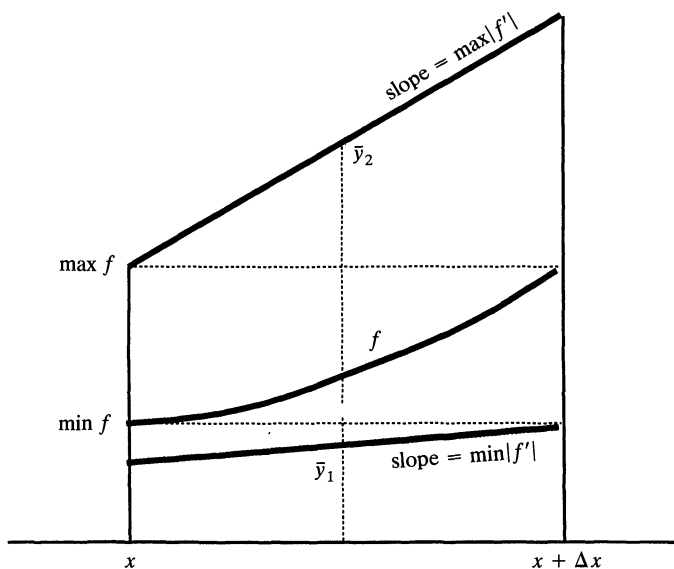


Figure 7

maximum value $\min_{[x, x+\Delta x]} f$ and slope $\min_{[x, x+\Delta x]} |f'|$; the upper has minimum value $\max_{[x, x+\Delta x]} f$ and slope $\max_{[x, x+\Delta x]} |f'|$. When revolved about the x -axis, each segment generates a frustum of a cone. According to Axiom (S), the areas of these frusta are lower and upper bounds (resp.) for the area generated by revolving the graph of f . Now, the area of a frustum is equal to $2\pi \bar{r}l$, where \bar{r} is the average of the two extreme radii of revolution, and l is the slant height, the length of the segment being revolved. (Slit and unroll again.) Consequently,

$$2\pi \bar{y}_1 \left(\min_{[x, x+\Delta x]} \sqrt{1 + f'^2} \right) \Delta x \leq S_x^{x+\Delta x} \leq 2\pi \bar{y}_2 \left(\max_{[x, x+\Delta x]} \sqrt{1 + f'^2} \right) \Delta x, \quad (5)$$

where \bar{y}_1 and \bar{y}_2 are the average radii (FIGURE 6). Since $\min_{[x, x+\Delta x]} f$ and $\max_{[x, x+\Delta x]} f$ both approach $f(x)$ as $\Delta x \rightarrow 0$, \bar{y}_1 and \bar{y}_2 also approach $f(x)$. The inequalities (5) therefore assert (B'), with $\varphi(x) = 2\pi f(x)\sqrt{1 + f'(x)^2}$. By the general theorem,

$$S_a^b = 2\pi \int_a^b f(x) \sqrt{1 + f'(x)^2} dx. \quad (6)$$

8. EQUIVALENCE OF THE AXIOMS WITH THE FORMULAS. Are axioms (L), (LL), and (S) convincing? Although self-evidence cannot be legislated, it may nevertheless help to know that each one, in conjunction with (A), is *equivalent* to the corresponding formula, (2), (4), or (6). We have seen in each case that the axiom, together with (A), implies the formula. Conversely, each formula implies (A), by additivity of the integral; and, clearly, (2) \Rightarrow (L), (4) \Rightarrow (LL₀) \Rightarrow (LL), and (6) \Rightarrow (S₀) \Rightarrow (S).

I wish to thank the referee for an extremely careful reading of the manuscript and for several thoughtful suggestions in addition to the contribution cited in Example 5. I am also grateful to Dan Kalman for a number of helpful comments.

REFERENCES

1. Leonard Gillman and Robert H. McDowell, *Calculus*, W. W. Norton, New York, 1973.
2. Hans Hahn and Arthur Rosenthal, *Set Functions*, The University of New Mexico Press, Albuquerque, 1948.
3. Serge Lang, *A First Course in Calculus*, Addison Wesley, Reading, Mass., 1968.
4. Howard Levi, *Polynomials, Power Series, and Calculus*, Van Nostrand, Princeton, 1968.

*Department of Mathematics
University of Texas
Austin, TX 78712*

Versatile Coins

István Szalkai and Dan Velleman

Suppose you had a coin which, when flipped, came up heads with probability

$$p = \frac{3 + \sqrt{3}}{6}$$

and tails with probability

$$1 - p = \frac{3 - \sqrt{3}}{6}.$$

If you flipped the coin three times, the probability of getting either three heads or three tails would be

$$p^3 + (1 - p)^3 = \frac{9 + 5\sqrt{3}}{36} + \frac{9 - 5\sqrt{3}}{36} = \frac{1}{2}.$$

Thus, by flipping this coin three times you could simulate the behavior of a fair coin, and thus make a random choice between two alternatives, with each alternative being chosen with probability $1/2$. But note that if you flipped the coin twice, the probability of getting one head and one tail (in either order) would be $2p(1 - p) = 1/3$, so you could also use this coin to make a random choice between two alternatives, with one alternative being chosen with probability $1/3$, and the other with probability $2/3$. In a sense, this coin is more versatile than either a fair coin or a coin which comes up heads with probability $1/3$.

We will say that $(3 + \sqrt{3})/6$ *simulates* both $1/2$ and $1/3$. In general, if p and q are any two numbers between 0 and 1, we will say that p *simulates* q if, given a coin which comes up heads with probability p , we can, in a finite number of flips, simulate the behavior of a coin which comes up heads with probability q . More precisely, p simulates q if there is some positive integer n and some subset E of the 2^n possible outcomes of flipping a coin n times such that, if a coin which comes up heads with probability p is flipped n times, the probability that the resulting sequence of heads and tails is in the set E is q . Now the probability $P(E)$ that the sequence of heads and tails is in the set E is given by the formula

$$P(E) = \sum_{i=0}^n a_i p^i (1 - p)^{n-i},$$

where for each i , a_i is the number of elements of E which contain i heads and $n - i$ tails. Clearly we have $0 \leq a_i \leq \binom{n}{i}$, and any sequence of integers $\{a_i\}_{i=0}^n$ satisfying these inequalities will correspond to some set E . Thus, p simulates q iff there is a positive integer n , and integers a_i satisfying $0 \leq a_i \leq \binom{n}{i}$, such that

$$q = \sum_{i=0}^n a_i p^i (1 - p)^{n-i}.$$

There are a number of observations that we can make right away. Clearly the “simulates” relation is reflexive, and it is not hard to see that it is also transitive. In other words, it is a preorder. Every number between 0 and 1 simulates both 0 and 1, since any coin can be used to simulate the behavior of a coin which either always comes up heads or always comes up tails. It is also clear that in general p and $1 - p$ simulate each other, and it follows by transitivity that if p simulates q then both p and $1 - p$ simulate both q and $1 - q$. It is not hard to see that if p simulates q and r , then it simulates their product qr . More generally, if p simulates q , r , and s , then it simulates $qr + (1 - q)s$ (see FIGURE 1). In fact, the reader might want to verify that the set of numbers simulated by p is precisely the closure of the set $\{0, 1\}$ under the function $f(x, y) = px + (1 - p)y$.

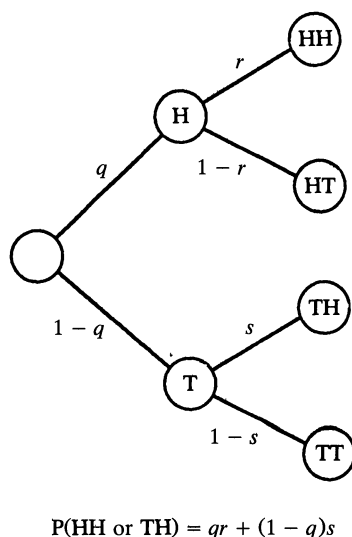


Figure 1. If we flip three coins, whose probabilities of coming up heads are q , r , and s , then we can find an event with probability $qr + (1 - q)s$. Thus, if p simulates q , r , and s , then it also simulates $qr + (1 - q)s$.

What else can we say about the “simulates” relation? One interesting way to investigate this relation is to try to design versatile coins, like the one described in the first paragraph of this paper. In that example we found that $(3 + \sqrt{3})/6$ simulates both $1/2$ and $1/3$. The reader may have been surprised that we used such a complicated number to simulate both $1/2$ and $1/3$. Couldn’t we have found a rational number that would do it?

The answer, it turns out, is no. The reason is that $1/2$ doesn’t simulate $1/3$, and no rational number other than $1/2$ simulates $1/2$. The first of these facts can be seen by noting that $1/2$ only simulates rational numbers with denominators of the form 2^n , for some natural number n . To prove the second, suppose $p = j/k$ is a rational number, with j and k relatively prime, and p simulates $1/2$. Then we can choose a positive integer n and integers a_i , $0 \leq a_i \leq \binom{n}{i}$, such that

$$\sum_{i=0}^n a_i p^i (1 - p)^{n-i} = \frac{1}{2}.$$

Now let $b_i = \binom{n}{i} - a_i$, and note that by the binomial theorem we have

$$\sum_{i=0}^n b_i p^i (1-p)^{n-i} = 1 - \frac{1}{2} = \frac{1}{2}.$$

Since $a_0 + b_0 = \binom{n}{0} = 1$, we have either $a_0 = 0$ or $b_0 = 0$. Without loss of generality, assume $a_0 = 0$. Then

$$\frac{1}{2} = \sum_{i=1}^n a_i p^i (1-p)^{n-i} = p \sum_{i=1}^n a_i p^{i-1} (1-p)^{n-i} = \frac{j}{k} \cdot \frac{\sum_{i=1}^n a_i j^{i-1} (k-j)^{n-i}}{k^{n-1}}.$$

Thus we have $k^n = 2j \cdot \sum_{i=1}^n a_i j^{i-1} (k-j)^{n-i}$, so $j|k^n$. Since j and k were assumed to be relatively prime, it follows that $j = 1$. But now note that $1-p = (k-j)/k = (k-1)/k$ also simulates $1/2$, so by the same reasoning we have $k-1 = 1$. Therefore $k = 2$, so $p = 1/2$. A similar, although slightly more complicated, proof can be used to show that the only rational numbers that simulate $1/3$ are $1/3$ and $2/3$. In fact, the proof can be generalized to show that for any square-free integer $N > 1$, the only rational numbers that simulate $1/N$ are $1/N$ and $(N-1)/N$.

Thus, we see that there are strict limits to the versatility of coins whose probability of coming up heads is rational. However, as the following theorem shows, coins with an irrational probability of coming up heads can sometimes be quite versatile.

Theorem 1. *Suppose F is a finite set of rational numbers and $F \subseteq [0, 1]$. Then there is a number $p \in [0, 1]$ such that p simulates every element of F .*

In the proof of the theorem, we will need the following lemmas:

Lemma 2. *For every integer $n > 1$,*

$$\left(1 - \frac{1}{n}\right)^{n-1} > \frac{1}{e}.$$

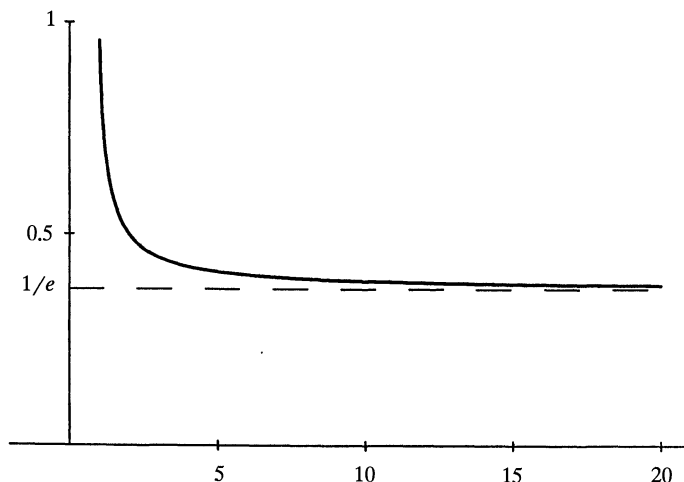


Figure 2. The Graph of the Function $f(x) = (1 - 1/x)^{x-1}$.

Proof: Let $f(x) = (1 - 1/x)^{x-1}$ for $x > 1$. It is clear that $\lim_{x \rightarrow \infty} f(x) = 1/e$, so to prove the lemma it will suffice to show that f is decreasing. We will let the reader check that

$$f'(x) = f(x) \left[\frac{1}{x} + \ln \left(1 - \frac{1}{x} \right) \right].$$

Since $\ln x < x - 1$ for all $x \in (0, 1)$, we have $\ln(1 - 1/x) < -1/x$ for all $x > 1$. Thus $f'(x) < 0$, so f is decreasing. \square

Lemma 3. *For every real number $z \in [0, 1/e]$ and positive integer n , there is a number $p \in [0, 1]$ such that*

$$np(1 - p)^{n-1} = z.$$

(Note that this equation says that if a coin which comes up heads with probability p is flipped n times, then the probability of getting exactly one head is z .)

Proof: If $n = 1$ then we simply let $p = z$. Now suppose $n > 1$. If $p = 0$ then the left side of the equation above is $0 \leq z$. If $p = 1/n$ then by Lemma 2 the left side is $(1 - 1/n)^{n-1} > 1/e \geq z$. Thus there is some $p \in [0, 1/n)$ such that $np(1 - p)^{n-1} = z$. \square

Proof of Theorem 1. Let N be the maximum of the denominators of the elements of F . We assume w.l.o.g. that $N \geq 4$. Let $n = N!/3$. By Lemma 3, let p be a solution to the equation $np(1 - p)^{n-1} = 1/3$. Then according to the analysis of the “simulates” relation above, p will simulate every number of the form

$$ap(1 - p)^{n-1} = \frac{a}{3n} = \frac{a}{N!},$$

for $0 \leq a \leq \binom{n}{1} = n = N!/3$. It follows that p simulates $1/3, 1/4, \dots, 1/N$.

As we observed above, if p simulates q then it also simulates $1 - q$, and if p simulates both q and r then it simulates their product qr . Thus, p simulates $2/3, 3/4, \dots, (N - 1)/N$. Since p simulates both $2/3$ and $3/4$, it also simulates $2/3 \cdot 3/4 = 1/2$. (This is where we use the assumption $N \geq 4$.) Now for any rational number j/k , with $1 \leq j < k \leq N$, we have

$$\frac{j}{k} = \frac{j}{j+1} \cdot \frac{j+1}{j+2} \cdot \dots \cdot \frac{k-1}{k}.$$

Since we have already shown that p simulates every factor on the right side, it follows that p simulates j/k . Thus p simulates every rational number between 0 and 1 with denominator at most N , and therefore in particular it simulates every element of F . \square

The same method of proof can also handle some sets of non-rational probabilities. Let F be a finite subset of $[0, 1/e]$ such that the ratio of any two nonzero elements of F is rational. Let z be the largest element of F . Then every other element of F can be written as a rational multiple of z , and by finding a common denominator for these rational multiples we may write F as

$$F = \left\{ z, \frac{zj_1}{n}, \frac{zj_2}{n}, \dots, \frac{zj_m}{n} \right\},$$

for some integers j_1, j_2, \dots, j_m and n with $0 \leq j_i < n$ for $1 \leq i \leq m$. Since $z \leq 1/e$, by Lemma 3 there is a number $p \in [0, 1]$ such that $np(1 - p)^{n-1} = z$. Then as

above p simulates every number of the form $ap(1-p)^{n-1} = za/n$, for $0 \leq a \leq n$. Since every element of F has this form, p simulates every element of F .

We can eliminate the upper bound $1/e$ in this result by using a somewhat more complicated proof. Instead of considering only sequences of coin flips in which there is *exactly* one head, we consider sequences which have *at least* one head.

Theorem 4. *Suppose $F \subseteq [0, 1]$, F is finite, and the ratio of any two nonzero elements of F is rational. Then there is a number $p \in [0, 1]$ such that p simulates every element of F .*

Proof: Since every number between 0 and 1 simulates 1, we may assume w.l.o.g. that $1 \notin F$. As before, we let z be the largest element of F , and then write F as

$$F = \left\{ z, \frac{zj_1}{N}, \frac{zj_2}{N}, \dots, \frac{zj_m}{N} \right\},$$

for some integers j_1, j_2, \dots, j_m and N with $0 \leq j_i < N$ for $1 \leq i \leq m$. Since $z < 1$, we can choose an integer n large enough so that

$$1 - \frac{1 + nN}{2^n} > z.$$

For each integer i , $1 \leq i \leq n$, let q_i be the quotient when $\binom{n}{i}$ is divided by N , and let r_i be the remainder. Thus q_i and r_i are nonnegative integers, $r_i < N$, and

$$\binom{n}{i} = Nq_i + r_i.$$

We now claim that we can choose $p \in [0, 1]$ so that p is a solution to the equation

$$(*) \quad \sum_{i=1}^n Nq_i p^i (1-p)^{n-i} = z.$$

Once we have such a p , we can conclude that for every integer a , $0 \leq a \leq N$, p simulates

$$\sum_{i=1}^n aq_i p^i (1-p)^{n-i} = \frac{za}{N}.$$

Since every element of F has this form, it follows that p simulates every element of F .

To see that we can choose $p \in [0, 1]$ satisfying $(*)$, first note that when $p = 0$, the left side of $(*)$ is $0 \leq z$. But when $p = 1/2$, the left side is

$$\begin{aligned} \sum_{i=1}^n \frac{Nq_i}{2^n} &= \frac{1}{2^n} \sum_{i=1}^n \left[\binom{n}{i} - r_i \right] = \frac{1}{2^n} \left[2^n - 1 - \sum_{i=1}^n r_i \right] \\ &> \frac{1}{2^n} [2^n - 1 - nN] = 1 - \frac{1 + nN}{2^n} > z. \end{aligned}$$

Thus, there is a value of p between 0 and $1/2$ which satisfies $(*)$. □

Note that in the proof of Theorem 1 above, p was chosen as a root of a polynomial with rational coefficients, so p was algebraic. In fact, in hindsight it is easy to see that this had to be true. If p simulates $q \in (0, 1)$, then by our characterization of the “simulates” relation there is a nonconstant polynomial $f(x)$ with integer coefficients such that $f(p) = q$. Thus if either p or q is algebraic then the other must be as well. More generally, we can say that $\mathbb{Q}[p] = \mathbb{Q}[q]$,

where for any real number a , $\mathbb{Q}[a] = \{r \in \mathbb{R} | r \text{ is algebraic over } \mathbb{Q}(a)\}$. This shows that if p simulates every element of some set $F \subseteq [0, 1]$, then for every $q \in F \setminus \{0, 1\}$, $\mathbb{Q}[q] = \mathbb{Q}[p]$. Thus, if $F \subseteq [0, 1]$ then there cannot be a number p which simulates every element of F unless for every $q, r \in F \setminus \{0, 1\}$, $\mathbb{Q}[q] = \mathbb{Q}[r]$.

For which sets $F \subseteq [0, 1]$ is there a number $p \in [0, 1]$ such that p simulates every element of F ? Our theorems so far do not completely settle this question. Although we do not know the complete answer, we can give the answer for the case $F \subseteq \mathbb{Q}$. We will need the following notation. For every positive integer N , let \mathbb{Q}_N be the set

$$\mathbb{Q}_N = \left\{ \frac{j}{N^k} \mid j, k \in \mathbb{Z} \right\}.$$

Our characterization of those sets $F \subseteq \mathbb{Q} \cap [0, 1]$ such that some number $p \in [0, 1]$ simulates all elements of F will be a consequence of the next two theorems.

Theorem 5. *Suppose $p \in [0, 1]$. Then there is some positive integer N such that $\{q \in \mathbb{Q} \cap [0, 1] | p \text{ simulates } q\} \subseteq \mathbb{Q}_N$.*

Proof: This proof is a modified version of a proof suggested to us by Martin Goldstern.

If p is not algebraic then, as we observed above, $\{q \in \mathbb{Q} \cap [0, 1] | p \text{ simulates } q\} = \{0, 1\}$, so the conclusion clearly holds. Now suppose p is algebraic. Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a minimal degree nonconstant polynomial with integer coefficients such that $f(p) = 0$. By multiplying through by -1 if necessary, we may assume that $a_n > 0$. We will show that $\{q \in \mathbb{Q} \cap [0, 1] | p \text{ simulates } q\} \subseteq \mathbb{Q}_{a_n}$.

Clearly $0, 1 \in \mathbb{Q}_{a_n}$. Now suppose $q \in \mathbb{Q} \cap (0, 1)$ and p simulates q . Then there is a nonconstant polynomial $g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0$ with integer coefficients such that $g(p) = q$. Thus $g(p) - q = 0$, so by the minimality of the degree of $f(x)$, $f(x)$ must divide $g(x) - q$. In other words, $g(x) - q = f(x) \cdot h(x)$ for some polynomial $h(x) = c_k x^k + c_{k-1} x^{k-1} + \cdots + c_1 x + c_0$ with rational coefficients. Multiplying out the product $f(x) \cdot h(x)$ and equating the coefficients with the coefficients of $g(x) - q$, we get the following equations:

$$\begin{aligned} m &= n + k \\ b_m &= a_n \cdot c_k \\ b_{m-1} &= a_n \cdot c_{k-1} + a_{n-1} \cdot c_k \\ b_{m-2} &= a_n \cdot c_{k-2} + a_{n-1} \cdot c_{k-1} + a_{n-2} \cdot c_k \\ &\vdots \\ b_1 &= a_1 \cdot c_0 + a_0 \cdot c_1 \\ b_0 - q &= a_0 \cdot c_0. \end{aligned}$$

Claim. For $i = 0, 1, \dots, k$, $(a_n)^{i+1} \cdot c_{k-i}$ is an integer.

Proof: By induction on i . The case $i = 0$ is taken care of by the second equation above, since b_m is an integer. For the general case we use the equation

$$b_{m-i} = a_n \cdot c_{k-i} + a_{n-1} \cdot c_{k-i+1} + \cdots.$$

Multiplying both sides by $(a_n)^i$ we get:

$$(a_n)^i \cdot b_{m-i} = (a_n)^{i+1} \cdot c_{k-i} + a_{n-1} \cdot (a_n)^i \cdot c_{k-i+1} + \cdots.$$

Clearly the left side of this equation is an integer, and by inductive hypothesis all

terms on the right side except the first are integers. Thus the first term must be an integer too. This proves the claim.

Applying the claim in the case $i = k$ we find that $(a_n)^{k+1} \cdot c_0$ is an integer. Let $j = (a_n)^{k+1} \cdot c_0$, so $c_0 = j/(a_n)^{k+1}$. Substituting this into the equation $b_0 - q = a_0 \cdot c_0$ we find that $q = b_0 - a_0 \cdot j/(a_n)^{k+1} \in \mathbb{Q}_{a_n}$, as required. \square

Theorem 6. *For every positive integer N there is a number $p \in [0, 1]$ such that p simulates every element of $\mathbb{Q}_N \cap [0, 1]$.*

Proof: Let N be any positive integer, and by Theorem 1 choose $p \in [0, 1]$ such that p simulates $1/2, 1/3, \dots, 1/N, 2/N, \dots, (N-1)/N$. To complete the proof we will show that for all positive integers j and k with $j < N^k$, p simulates j/N^k .

We proceed by induction on k . The case $k = 1$ is clear, by the choice of p . For the induction step, suppose that for every positive integer $j < N^k$, p simulates j/N^k , and let j be any positive integer less than N^{k+1} . We must show that p simulates j/N^{k+1} .

Let q and r be the quotient and remainder when j is divided by N^k . Then $j = qN^k + r$, $0 \leq q < N$, and $0 \leq r < N^k$. By the choice of p , p simulates q/N and $1/(N-q)$, and by inductive hypothesis p simulates r/N^k . Now we apply the fact, observed above, that if p simulates x , y , and z , then it simulates $xy + (1-x)z$. Taking $x = q/N$, $y = 1$, and $z = 1/(N-q) \cdot r/N^k$, we see that p simulates

$$\frac{q}{N} + \left[1 - \frac{q}{N}\right] \frac{r}{(N-q)N^k} = \frac{q}{N} + \frac{r}{N^{k+1}} = \frac{qN^k + r}{N^{k+1}} = \frac{j}{N^{k+1}},$$

as required. \square

Combining Theorems 5 and 6, we get the following characterization of those sets of rational probabilities which can be simulated by a single number.

Corollary 7. *Suppose $F \subseteq \mathbb{Q} \cap [0, 1]$. Then there is a number $p \in [0, 1]$ such that p simulates every element of F iff for some positive integer N , $F \subseteq \mathbb{Q}_N$.* \square

By examining the proofs of some of the theorems above, we can actually determine precisely what rational probabilities are simulated by the numbers used in some of our proofs. Suppose N is an integer, $N \geq 4$, and let p be a solution to the equation $np(1-p)^{n-1} = 1/3$, where $n = N!/3$. As we showed in the proof of Theorem 1, p simulates all rational probabilities with denominator at most N . In fact, this also holds if $N = 3$, as the example in the first paragraph of this paper shows. The proof of Theorem 6 now shows that p also simulates all elements of $\mathbb{Q}_N \cap [0, 1]$. In fact, we can improve on the argument in that proof to show that p simulates all elements of $\mathbb{Q}_{N!} \cap [0, 1]$. This is easily seen to follow from the following proposition.

Proposition 8. *Let p and N be as described above. Suppose M and k are positive integers, $2 \leq k \leq N$, and p simulates all rational numbers of the form j/M , for $0 < j < M$. Then p simulates all rational numbers of the form $j/(Mk)$, for $0 < j < Mk$.*

Proof: Suppose $0 < j < Mk$. Let q and r be the quotient and remainder when j is divided by M . Then $j = Mq + r$, $0 \leq q < k$, and $0 \leq r < M$. Thus p simulates

q/k , $1/(k - q)$, and r/M . As in the proof of Theorem 6, it follows that p simulates

$$\frac{q}{k} + \left[1 - \frac{q}{k}\right] \cdot \frac{1}{k - q} \cdot \frac{r}{M} = \frac{Mq + r}{Mk} = \frac{j}{Mk}. \quad \square$$

We can learn more about the rational numbers simulated by p by examining the proof of Theorem 5. Clearly p is algebraic, since it was chosen to be a root of the polynomial $g(x) = 3nx(1 - x)^{n-1} - 1 = (N!)x(1 - x)^{n-1} - 1$. Let $f(x)$ be a minimal degree nonconstant polynomial with integer coefficients such that $f(p) = 0$, and let a be the coefficient of the highest power of x in $f(x)$. The proof of Theorem 5 shows that every rational probability simulated by p must be in the set \mathbb{Q}_a . What can we say about the value of a ?

Recall that the *content* of a polynomial with integer coefficients is defined to be the greatest common divisor of its coefficients. A polynomial is called *primitive* if its content is 1, and Gauss' Lemma says that the product of two primitive polynomials is also primitive. Note that $g(x)$ is primitive, since its constant term is -1 , and w.l.o.g. we may assume that $f(x)$ is primitive as well.

By the minimality of the degree of $f(x)$, $g(x)$ is divisible by $f(x)$, so $g(x) = f(x) \cdot h(x)$, for some polynomial $h(x)$ with rational coefficients. By finding a common denominator for the coefficients of $h(x)$, and then factoring out the greatest common divisor of their numerators, we may write $h(x) = (j/k) \cdot h'(x)$, for some primitive polynomial $h'(x)$. Thus $k \cdot g(x) = j \cdot f(x) \cdot h'(x)$. But now the content of the left side of this equation is k , and by Gauss' Lemma the content of the right is j , so $j = k$, and therefore $h(x) = h'(x)$, so $h(x)$ is primitive.

Since the coefficient of the highest power of x in $g(x)$ is $\pm N!$ and $g(x) = f(x) \cdot h(x)$, it follows that $a|N!$, and therefore $\mathbb{Q}_a \subseteq \mathbb{Q}_{N!}$. Combining this with our earlier conclusions that p simulates all elements of $\mathbb{Q}_{N!} \cap [0, 1]$, and that all rational probabilities simulated by p are in \mathbb{Q}_a , we see that

$$\{q \in \mathbb{Q} \cap [0, 1] \mid p \text{ simulates } q\} = \mathbb{Q}_{N!} \cap [0, 1].$$

For example, in the case $N = 3$ we can conclude that the set of rational probabilities simulated by the number $(3 + \sqrt{3})/6$ is precisely $\mathbb{Q}_6 \cap [0, 1]$.

There are still many unanswered questions about the simulation of irrational probabilities. One of the simplest is this: If q and r are algebraic numbers between 0 and 1, must there be a number $p \in [0, 1]$ such that p simulates both q and r ?

*Department of Mathematics
University of Veszprém
H 8201 Veszprém
Hungary
h2109sza@ella.hu*

*Dept. of Mathematics and Computer Science
Amherst College
Amherst, MA 01002
djvellenman@amherst.edu*

Two-Year Magazine Subscription Rates

Underwood Dudley

When a magazine incites you to subscribe to it, you are often given the choice of signing up for one year or two. The cost of the second year is usually less than that of the first. How do magazines decide on how much the discount for the second year should be? Is a mathematical model at work?

A simple model could take into account the interest rate, i , that the magazine could earn on money paid in advance, the cost, c , of obtaining a subscription, and the probability, p , that a one-year subscriber will renew. If we let s denote the single-year subscription rate and t the two-year rate, then the present value of a two-year subscription is $t - c$ while the present value of two successive one-year subscriptions is

$$s - c + \frac{p(s - c)}{1 + i}.$$

The magazine will neither gain nor lose if the two amounts are equal, which gives

$$t = s \left(1 + \frac{p}{1 + i} - \frac{(c/s)p}{1 + i} \right). \quad (1)$$

So, for example, if $p = .7$, $i = .1$, and $c/s = .2$, then $t = 1.50s$. That is, the magazine can give a 50% discount for the second year's subscription and not lose money. For a prosperous magazine with loyal readers, say $p = .95$, $i = .05$, and $c/s = .1$, we get $t = 1.90s$, a 10% discount, while for a magazine with a high turnover of readers ($p = .5$), forced to use new funds to pay off high-interest debt ($i = .15$), and with large promotion costs ($c/s = .4$), the discount would approach 75%: $t = 1.26s$.

Of course, if the magazine could give a smaller discount and attract the same number of subscribers, the extra income would be sheer gain. Since subscribers do not carry out the same calculations as the magazine, this may be possible. The variables of interest to the subscriber are j , the interest rate the subscriber could earn on magazine-subscription money and q , the probability of disenchantment with the magazine before the end of a year's subscription. The present value to the subscriber of a second year's subscription, discounted for interest and disenchantment, is $s(1 - q)/(1 + j)$, so the amount a subscriber would pay for a two-year subscription is

$$t = s \left(1 + \frac{1 - q}{1 + j} \right).$$

Since $1 - q$ will probably be greater than p since people contemplating buying two-year subscriptions are less likely to become disenchanted with a magazine than are the mass of its subscribers, many of whom may be merely giving the magazine a

try, it follows that many subscribers would accept a smaller discount than (1) would indicate. For example, if $j = .05$ and $q = .2$, then $t = 1.76s$, and even if $j = .20$ (using money to subscribe that could have gone to reducing credit-card debt) and $q = .4$, then $t = 1.50s$.

Since few magazines have p as large as .95, the simple model suggests that discounts for second-year subscriptions should probably be somewhere in the range between 25% and 50%. The model, however, does not reflect reality. An unscientific sample of 39 magazines gave the following results:

% discount	0–9	10–14	15–19	20–24	25–29	30–39	40–49	50–
% of magazines	8	10	15	23	10	18	6	10.

The distribution of discounts is fairly uniform, with a median of 24% ($\bar{x} = 25.9$ and $s = 14.2$). Three magazines in the sample (*Paris Review*, *Harper’s Bazaar*, and *Country Living*) give no discount at all, completely disregarding their readers’ ability to make calculations of present values. Discounts are less than what the model would predict.

One reason for the discrepancy is that the model may be incorrect. It takes into account neither inflation nor the fact that a magazine’s subscription rate next year is this year’s rate plus a random variable with positive mean. A different reason is that magazines may be using a different model. This is true in a sense, since it turns out that most magazines use no model at all. I sent a letter to 65 magazines (again, unscientifically selected) asking how two-year subscription rates were determined, soliciting detailed information. For those who did not choose to go into detail, I included a sheet with four choices that could be checked off. The 32 returns (49% of those surveyed) were distributed as follows:

Choice	Number	Percent
Rates are determined by taking into account renewal rates and interest rates	2	7
Rates are determined by judgment and experience, without technical analysis	26	90
It’s none of your business how we set our rates	0	0
Go away, we’re busy	1	3

Three respondents declined to choose any of the above. One wrote, “Mostly by postal rates. When USPS goes up, so do we.” Another responded, in full, “We use the Magick 8 Ball. No joke.”

Of the two respondents who marked the first choice, one crossed off *and interest rates*. The other added a marginal note, “a formula” but gave no details. Since this response also included “Go Wabash!” it was not clear how much weight it should be given.

The inescapable conclusion is that mathematics does not enter into the rate-setting process. Some of the responses indicated what I interpreted as lack of

mathematical sophistication:

We want to show there is an advantage of taking a 2 year offer, so we chose to reduce the price of two years at the 1 year offer price by \$2. I. e., 1 year \$15, 2 years \$28.

You multiply the one year rate by 2, 3, etc. and give a savings as an incentive.

One verged on the antimathematical:

Multiple year rates are then re-set, based on discounts off the one year rate. Decisions are almost always subjective, and do not take into account such factors as inflation, interest rates, etc. Circulation is a statistical enterprise, not mathematical!

Most, however, reflected empiricism:

We usually charge less for the second year. That works the best. Sometimes we test higher prices, but it's all based on judgment and past experience.

Rates are tested in direct mail panels to see which pull the best response but are still the most profitable.

We are constantly testing pricing, both up and down.

We look at the competition and what their rates are, the time of year, the economic forecasts, etc., etc. Then we make up our minds hoping we have guessed right.

One was faintly despairing:

We're a membership organization. Rates are set by delegates and voting members. No rhyme or reason.

Of course, there are many factors to be considered in setting rates. For example, the Audit Bureau of Circulation will not count in a magazine's paid circulation any subscription taken out at less than 50% of the base rate. Advertising, for many magazines as important as circulation, must be taken into account:

Subscription pricing also is a factor in advertising sales—some advertisers place a higher value on subscribers who pay more.

In some promotions a publisher will net (subscription revenue-promotion expense) as little as \$1. If they can opt for multi-year, the net-per-year increases greatly.

The facts are that only about 50% of the subscribers will renew. On the two-for-one offer, my 456,000 subscribers are there for two years to help me meet my rate base target. The one year offer will lose half its subscribers at the end of year one, so they must also be replaced (perhaps at a loss) and that must be considered along with the renewal profit made from the ones who do renew. The complexity expands rapidly, so you can see why we use a computer model to evaluate this stuff.

I think it is clear that most magazines do not seek the aid of mathematics when they come to set their two-year subscription rates. It is also clear to me, though it cannot be demonstrated from the survey results, that most magazines do not even *think* of using mathematics. Mathematics, as one respondent pointed out, is irrelevant.

How can that be? Almost any textbook you pick up, on subjects from intermediate algebra, through calculus, to linear algebra and beyond, has *With Applications* in its title and in its preface brags about the number, variety, and value of the applications to be found in it. Our colleges graduate annually hundreds of thousands of people who are certified as having mastered the contents of such texts, and some of them inevitably enter the employ of magazines. Yet an insignificant number of magazines use even the very simple mathematics that would be needed to construct a useful model of subscription revenue.

If mathematics is not being applied where it could easily be applied, my conclusion is that the emphasis on applications is certainly useless and perhaps harmful. When students eventually discover that they in fact never need or apply any of what they have been told is so applicable and necessary, disillusion with mathematics may result, word may get around, and enrollments may decline.

I think that it is better to present mathematics to students as a glorious adventure for the mind. (It is in fact the greatest achievement of the human mind, but I do not tell students that because they might mistakenly think me guilty of overstatement.) That it has uses is important, but incidental. Few students will use it, but all can see some of the glory.

*Department of Mathematics and Computer Science
DePauw University
Greencastle, IN 46135
dudley@depauw.bitnet*

“The defects in the mathematical training of the student of engineering appear to be largely in the knowledge and grasp of fundamental principles, and the constant effort of the teacher should be to ground the student thoroughly in these fundamentals, which are too often lost sight of in a mass of details.”

—*American Mathematical Monthly*
18, (1911) p. 24.

On Seeing Progressions of Constant Cross Ratio

R. J. Duffin

I. GEOMETRIC, ARTISTIC, AND MELODIC PROGRESSIONS. A geometric progression is characterized by the property that every two successive terms, say a and b , have a common ratio, $r = b/a$. We shall say that a progression is *artistic* if every three successive terms, say a , b , and c , have a common *cross ratio* P , here defined as

$$P[a, b, c] = \frac{(a + b + c)b}{ac} = \frac{(\text{whole length})(\text{middle})}{(\text{one end})(\text{other end})}.$$

All terms are interpreted as positive lengths. The concept of cross ratio was introduced in the third century by the Greek geometer Pappus [2].

Theorem 1.1. *A geometric progression with common ratio $r > 0$ is also an artistic progression. Its common cross ratio satisfies*

$$P = 1/r + 1 + r \geq 3.$$

Proof: If a, b, c are three successive terms of a geometric progression with common ratio r , then $b = ar$ and $c = br$. Therefore

$$P = \frac{(a + b + c)b}{ac} = \frac{(b/r + b + br)b}{bb} = 1/r + 1 + r.$$

The inequality $r + 1/r \geq 2$ gives the bound $P \geq 3$. However, an artistic progression may not be a geometric progression.

These concepts lead to a new property of a progression which we term the *bias* Q , defined as

$$Q[a, b, c] = \frac{ac - b^2}{abc}.$$

Theorem 1.2. *A progression is geometric if and only if it has zero bias.*

Theorem 1.3. *A progression is artistic if and only if it has a common bias.*

The proof of Theorem 1.2 is obvious. Theorem 1.3 is somewhat of a surprise; it is a corollary of Theorem 1.5 to follow.

Jacobi's dictum for mathematical insight is, "Always invert!" Here, a simple minded application of Jacobi's advice is to introduce the sequence of reciprocal lengths. But frequency is inversely proportional to wavelength. Thus, one can refer

to the terms of the new progression as frequencies. So let frequencies be denoted as

$$A = 1/a, B = 1/b, C = 1/c, \text{ etc.}$$

If a, b, c, d, \dots is an *artistic progression*, then we say that A, B, C, D, \dots is a *melodic progression*.

The introduction of this extra terminology simplifies the algebra. Moreover, melodic progressions have the surprising property of satisfying a linear recursion formula, similar to that satisfied by the Fibonacci progressions.

If A, B , and C are three successive terms in a progression of frequencies then P^* , the *frequency cross ratio*, and Q^* , the *frequency bias*, are defined as:

$$P^* = A/B + C/B + AC/B^2,$$

$$Q^* = B - AC/B.$$

It is clear that this definition gives $P^* = P$ and $Q^* = Q$ for the lengths.

The word “melodic” is used here only as an algebraic term. For example, here is a melodic progression with cross ratio 3 and bias 10:

$$30, 60, 100, 150, 210, 280, 360, 450, 550, 660, 780, 910, 1050.$$

Would this scale offend the musical ear? This could easily be tested with a computer.

Theorem 1.4. *If A, B, C are three successive terms in a melodic progression, the next term is*

$$D = AC^2/B^2 + C^2/B - C.$$

This is the general recursion formula for melodic progressions.

Proof: From the definition of melodic cross ratio it follows that:

$$P + 1 = \left(1 + \frac{A}{B}\right)\left(1 + \frac{C}{B}\right).$$

If a progression A, B, C, D is melodic, then

$$\left(1 + \frac{A}{B}\right)\left(1 + \frac{C}{B}\right) = \left(1 + \frac{B}{C}\right)\left(1 + \frac{D}{C}\right).$$

So $C^2(B + A) = B^2(C + D)$. Then solving for D yields the formula.

Theorem 1.5. *A progression is melodic if and only if the frequency bias,*

$$Q = B - AC/B$$

always has the same value.

Proof: It is sufficient to note that the relation

$$B - AC/B = C - BD/C,$$

when multiplied by C/B is equivalent to the identity of Theorem 1.4.

Theorem 1.6. *If an artistic progression has positive bias, then the progression is strictly convex. If an artistic progression has negative bias, then the corresponding melodic progression is strictly convex.*

Proof: If $Q > 0$, then $Q = (ac - b^2)/abc > 0$. Thus $b < \sqrt{ac}$. But by the geometric mean inequality $\sqrt{ac} \leq (a + c)/2$. Hence, $2b < (a + c)$ and $(b - a) < (c - b)$. Repeating the argument shows

$$(b - a) < (c - b) < (d - c) < \cdots.$$

This defines strict convexity for the artistic progression.

If $Q < 0$, then $Q = (B^2 - AC)/B < 0$. Thus $B < \sqrt{AC} \leq (A + C)/2$. Then the similarity with the previous case shows that

$$(B - A) < (C - B) < (D - C) < \cdots$$

This defines strict convexity for the melodic progression.

II. A LINEAR RECURSION FORMULA FOR MELODIC PROGRESSIONS. Now suppose both the cross ratio and the bias are specified.

Theorem 2.1. *A melodic progression has cross ratio P and bias Q . Then A and B are successive terms in this progression if and only if*

$$A^2 + B^2 + (1 - P)AB = Q(A + B).$$

Proof: We recognize two recursion formulas:

$$C = \frac{B(PB - A)}{A + B} \quad \text{and} \quad C = \frac{B(B - Q)}{A}.$$

Setting them equal gives

$$PAB - A^2 = (B - Q)(A + B)$$

which by rearranging establishes the stated equation.

Corollary 2.1. *If a melodic progression has $P = 3$ then*

$$Q = \frac{(A - B)^2}{A + B}.$$

Theorem 2.2. *Let A, B, C be three successive terms in a melodic progression having cross ratio P and bias Q . Then*

$$C = (P - 1)B - A + Q.$$

is a linear recursion formula for the progression.

Proof: The first recursion formula used in the proof of Theorem 2.1 gives

$$\begin{aligned} C(A + B) &= PB^2 - AB, \\ CB &= PB^2 - AB - CA, \\ CB &= (P - 1)B^2 - AB - CA + B^2. \end{aligned}$$

But by definition $QB = B^2 - CA$ so

$$CB = (P - 1)B^2 - AB + QB.$$

Dividing by B yields the stated recursion formula.

Theorem 2.3. *Let A, B, C, D, \dots be a melodic progression with cross ratio $P > 3$ and bias Q . Let Σ be the melodic series*

$$\Sigma = A + B + C + D + \cdots$$

Then the n th term of this series is

$$A_{n-1} = Hr^{n-1} + Jr^{1-n} + K.$$

Here

$$2r = (P - 1) - \sqrt{(P - 1)^2 - 4}.$$

$$K = Q/(P - 3),$$

$$(1/r - r)H = (A - K)/r - (B - K), \text{ and}$$

$$(1/r - r)J = -(A - K)r + (B - K).$$

Proof: The linear recursion formula for the melodic progression is of the form

$$C = (P - 1)B - A + Q.$$

In difference equation notation, this may be written as the difference equation

$$A_{m+2} - (P - 1)A_{m+1} + A_m = Q.$$

One solution of the homogeneous difference equation

$$A_{m+2} - (P - 1)A_{m+1} + A_m = 0$$

is $A_m = r^m$, where $r^2 - (P - 1)r + 1 = 0$. An independent solution is r^{-m} .

A particular solution of the difference equation is K itself. Thus, the general solution of the difference equation is of the form stated with H and J arbitrary constants. To satisfy the initial conditions for $m = 0$ and $m = 1$ we choose A and B as follows:

$$\begin{aligned} A &= H + J + K, \\ B &= Hr + J/r + K. \end{aligned}$$

Solving these equations for H and J gives the stated expressions and completes the proof.

Corollary 2.2. *The sum of n terms of the progression of Theorem 2.3 is*

$$\sum_n = H \frac{1 - r^n}{1 - r} + J \frac{1 - r^{-n}}{1 - r^{-1}} + Kn.$$

Proof: The derivation of this relation follows directly from the sum of three different geometric progressions.

The following two theorems are stated without proof. They are corollaries of Theorems 2.1 and 2.2.

Theorem 2.4. *Suppose that A, B, C are three successive terms in a melodic progression. Then:*

$$2(C \text{ or } A) = (PB - B - Q) + \sqrt{[(PB - B - Q)^2 + 4(QB - B^2)]},$$

$$2(A \text{ or } C) = (PB - B - Q) - \sqrt{[(PB - B - Q)^2 + 4(QB - B^2)]}.$$

This gives a one term recursion formula for melodic progressions.

Theorem 2.5. *If A, B, C, D are four successive terms in a melodic progression and $C \neq B$, then the melodic cross ratio is*

$$P = \frac{D - A}{C - B}.$$

III. A KEY THEOREM APPLYING TO VISION, DRAWINGS, AND PHOTOGRAPHS. A line segment partitioned into three parts is termed a *triad*. Suppose a is the length of the first part, b is the length of the second part and c is the length of the third part. The triad is symbolized as $[a, b, c]$.

If an object is a triad, then its photographic image is a triad. To analyze the positive film (not the negative film) we may use FIGURE 1, following a standard method used to explain drawing. The object line and the image line are called transversals of the rays to the eye.

Then objects and their photographic images are subject to the following theorem of Pappus.

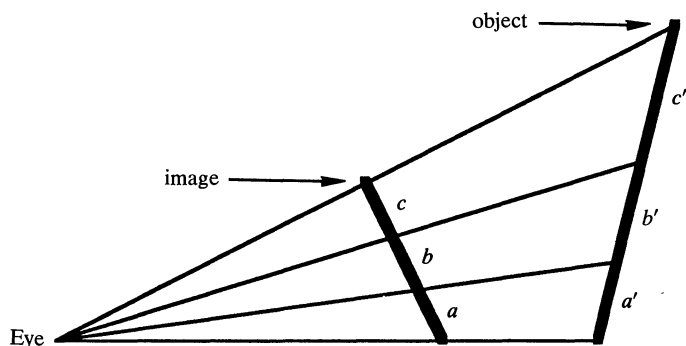


Figure 1. An image of a triad.

Theorem P. *Cross ratios of image triad $[a, b, c]$ and object triad $[a', b', c']$ are equal.*
Thus

$$\frac{(a + b + c)b}{ac} = \frac{(a' + b' + c')b'}{a'c'}.$$

For example, consider the following puzzle: A freehand artist draws an image of a yardstick on a mural. He reduces the first foot of the yardstick to 21 centimeters and the second foot to 14 centimeters. How long should he make the image of the third foot? The answer to this puzzle is an integer. Many people guess 7 centimeters.

To get the correct answer the artist simply solves the P equation for c , and obtains

$$c = \frac{b(a + b)}{Pa - b}.$$

In the case of a yardstick, $a' = b' = c'$ so $P = 3$ and

$$c = \frac{b(a + b)}{3a - b}.$$

Let $a = 21$ cm and $b = 14$ cm. So the image of the third foot is 10 cm.

All photographic images of yardsticks have cross ratio 3. Moreover, by applying Theorem P twice, all photographs of photographs of yardsticks have cross ratio 3. Thus, one cross ratio has a remarkable indestructible property because a photo-

graph of a photograph is a counterfeit and may not be congruent to a photograph of the real world.

Triads with cross ratio 3 are so common that they may be called *normal triads*.

IV. ONE SEES BRICKS ALMOST EVERY DAY. Bricks of uniform size are ubiquitous in present day civilization and in ancient civilizations as well. Consider a photograph of a brick wall. Every three bricks form a normal triad with cross ratio 3. Now consider four bricks in a row. If the first has image 21 mm, and the second has image 14 mm, then we have seen that the third image must be 10 mm. We obtain the length of the fourth image using $c = b(a + b)/(3a - b)$ with $a = 14$, $b = 10$. We find $c = 7.5$ mm. The total length of the image is $21 + 14 + 10 + 7.5 = 52.5$ mm.

A general algorithm for such problems is stated as follows:

Theorem 3.1. *Let a_1, a_2, a_3, \dots be image lengths of bricks aligned in a long wall. If $a_1 > a_2$, then the sum of $n > 2$ lengths, $a_1 + a_2 + a_3 + \dots + a_n$, is*

$$S_n = \frac{nS}{T + n}$$

where if $a_1 = a$ and $a_2 = b$ then

$$S = \frac{a(a + b)}{a - b} \quad \text{and} \quad T = \frac{2b}{a - b}.$$

The n th term is

$$a_n = \frac{ST}{(T + n)(T + n - 1)}.$$

The infinite sum is S .

Proof: Let a line of n bricks of unit length be partitioned as a triad with the first brick the first part of the triad, the second brick the second part of the triad, and the remaining $n - 2$ bricks the third part of the triad. The cross ratio of the object triad is $P[1, 1, n - 2]$. The cross ratio of the image triad is $P[a, b, S_n - a - b]$. So by Theorem P

$$P[a, b, S_n - a - b] = P[1, 1, n - 2].$$

Evaluating each side of this equation gives

$$\frac{bS_n}{a(S_n - a - b)} = \frac{n}{n - 2}.$$

Solving for S_n proves that

$$S_n = \frac{na(a + b)}{(a - b)n + 2b} = \frac{nS}{T + n}.$$

Since $a_n = S_n - S_{n-1}$ we obtain

$$a_n = \frac{nS}{T + n} - \frac{(n - 1)S}{T + n - 1} = \frac{ST}{(T + n)(T + n - 1)}.$$

The last assertion of the theorem follows by letting $n \rightarrow \infty$ in S_n .

However, suppose that an eccentric architect designs a brick wall with the n th brick of length r^{n-1} . The length of n such bricks is

$$G_n = 1 + r + r^2 + \cdots + r^{n-1} = (1 - r^n)/(1 - r).$$

Theorem 3.2. *Let a series*

$$S = a + b + c + d + \cdots$$

be the termwise image of a geometric series with ratio $r > 0$. If $a > br$ and $r \neq 1$, then S is an artistic series with cross ratio

$$P = 1/r + 1 + r.$$

The sum of $n > 2$ terms is

$$S_n = \frac{a(a+b)}{a - b(r - r^{n-1})/(1 - r^n)}.$$

The infinite sum of image lengths is

$$S = \frac{a(a+b)}{a - br}.$$

Proof: We may partition the architect's n bricks as a triad, as in the proof Theorem 3.1. Then the stated relations follow from Theorem P .

Corollary 3.1. *Let $S = a + b + c + d + \cdots$ be an artistic series having cross ratio $P > 3$ and $a > br$. Here r is defined by*

$$2r = (P - 1) - \sqrt{(P - 1)^2 - 4}.$$

Then the sum of n terms of the series is

$$S_n = \frac{a(a+b)}{a - b(r - r^{n-1})/(1 - r^n)}.$$

Proof: It is sufficient to note that r so defined, is a root of the equation

$$P = 1/r + 1 + r$$

and that an artistic series is determined by P and its first two terms.

V. MELODIC INTEGER PROGRESSIONS. If the coefficients of a linear recursion formula are integers and the initial conditions are integers, then an infinite sequence of integers is generated. The proof of the following theorem depends on this property.

Theorem 4.1. *Let x and y be arbitrary parameters then a unique melodic progression A, B, C, \cdots results by the assignment:*

$$\begin{aligned} A &= 1, \\ B &= x > 0, \\ P &= xy + y - 1, \\ Q &= -xy + x + 1. \end{aligned}$$

Moreover, if x and y are integers, then all terms of the progression are integer valued.

Proof: According to Theorem 2.1 a melodic progression results if

$$A^2 + B^2 + (1 - P)AB = Q(A + B).$$

The given assignment of values for A, B, P, Q satisfy this equation. Thus a melodic progression (or regression) results with the recursion formula

$$C + A = (P - 1)B + Q.$$

If x and y are integers, then P and Q are integers. Thus the progression is integer valued, extending forward without end. The coefficient of A is unity and this implies that the regression is integer valued, extending backward without end.

Here are some examples of melodic integer progressions obtained by choosing parameters x and y in the above algorithm:

[1] $P = 3, Q = 1$ gives

3, 1, 0, 0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105,

[2] $P = 3, Q = 9$ gives

119, 77, 44, 20, 5, -1, 2, 14, 35, 65, 104, 152, 209, 275, 305,

[3] $P = 5, Q = 1$ gives

1, 0, 0, 1, 5, 20, 76, 285, 1065, 3976, 14840, 55385,

[4] $P = 5, Q = -1$ gives

2, 1, 1, 2, 6, 21, 77, 286, 1066, 3977, 14841, 55386,

[5] $P = 11, Q = -5$ gives

2291, 232, 24, 3, 1, 2, 14, 133, 1311, 12972, 128404, 1271063.

The progression [1] is the well-known sequence of triangular numbers. Possibly there may be other diophantine relations arising. For example, the integer progressions [3] and [4] were discovered to have terms differing by unity. Is that a rare property?

The progressions were evaluated by the linear recursion formula, but note that uninvited negative and zero terms stole in. How is the theory extended if the ratios r and P are allowed to be complex or even quaternionic [2]?

VI. GEOMETRICAL EVALUATION OF THE BIAS OF A BRICK LINE. Here we compute the numerical value of the bias, given the object line, the eye point and the image line.

Theorem 5.1. *The bias of the image of a line of bricks is*

$$Q = \frac{2j(\sin \phi)^2}{UV}.$$

Here: (1) j is the length of a brick, (2) U is the length of the normal from the eye to the brick line, (3) V is the length of the normal from the eye to the image line, (4) ϕ is the angle between these normals.

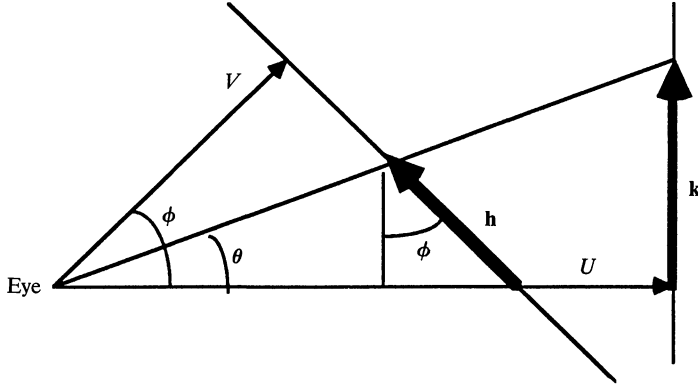


Figure 2. The image h of a line k .

Proof: In the large right triangle, $\cot \theta = U/k$. In the small one,

$$\cot \theta = [V/\cos \phi - h \sin \phi]/h \cos \phi = U/k.$$

Rearranging terms in this identity gives

$$1/h = M/k + N$$

where $M = (\cos \phi)^2 U/V$ and $N = (\sin \phi \cos \phi)/V$.

Adding one brick to the brick line increases the lines length by j and increases the length of the image by an amount a . Thus

$$h = k/[M + Nk] \quad \text{and} \quad h + a = [k + j]/[M + N(k + j)].$$

Adding another brick causes the image to increase by an amount b

$$h + a + b = [k + 2j]/[M + N(k + 2j)].$$

By subtraction, these expressions give values for a and b :

$$a = Mj/[M + N(k + j)][M + Nk] \quad \text{and} \\ b = Mj/[M + N(k + 2j)][M + N(k + j)].$$

Let $A = 1/a$ and $B = 1/b$. Thus

$$A = [M + N(k + j)][M + Nk]/Mj \quad \text{and} \\ B = [M + N(k + 2j)][M + N(k + j)]/Mj.$$

To use the expression for Q given by Corollary 2.1,

$$Q = (B - A)^2/(B + A),$$

it is found that

$$B - A = 2[M + N(k + j)]N/M \quad \text{and} \\ B + A = 2[M + N(k + j)]^2/Mj.$$

Substituting in the Corollary 2.1

$$Q = 2jN^2/M = 2j[(\sin \phi \cos \phi)/V]^2/[(\cos \phi)^2 U/V].$$

This yields the stated value of Q and completes the proof.

A consequence of Theorem 5.1 is that the bias is independent of a shift of the bricks along their line. In other words:

Corollary 5.1. *The bias of an image of a train of cars is constant even though the train is moving along its track.*

Books explaining the theory of relativity often have examples involving moving trains. Possibly Corollary 5.1 would not hold at half the speed of light because there are distortions due to fast motion, such as the Fitzgerald contraction [3].

VII. DISCUSSION. The above developments started from a simple puzzle about drawing a yardstick. One thing led to another and many different mathematical concepts were related. Several questions have been left unanswered. No doubt, readers will see other questions too.

REFERENCES

1. Howard Eves, *An Introduction to the History of Mathematics*, page 163, Holt, Rinehart and Winston, 1969.
2. Richard Duffin, A generalization of the ratio test for series, *Amer. Math. Monthly* 55, (1948) 153–155.
3. Ping-Kang Hsiung, Robert H. Thibadeau, Michael Wu, T-buffer: fast visualization of relativistic effects in spacetime, *ACM Trans.* (1990) 83–88.

*Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA 15213*

Report on the First Putnam Competition

The five persons ranking highest in the examination, arranged alphabetically, were ROBERT GIBSON, Fort Hays Kansas State College; I. KAPLANSKY, University of Toronto; G. W. MACKEY, Rice Institute; M. J. NORRIS, College of St. Thomas, BERNARD SHERMAN, Brooklyn College. Each of these will receive a prize of fifty dollars. The order of the names in this list has no relation to their rank in the examination.

—*American Mathematical Monthly*
45, (1938) p. 332.

100 Years of Monthly Editors

The *Monthly* was founded in 1894 by B. F. Finkel. In 1913 it was reorganized under an editorial board consisting of B. F. Finkel and representatives from eleven supporting colleges. By 1914 the number of those institutions had grown to fourteen. Below are listed the boards of editors up to 1916, with the year they began serving. In 1913 H. E. Slaught became Managing Editor and continued to hold that position until 1918 when the post was renamed Editor-in-Chief. In 1916 the *Monthly* became the Association's official journal. From that date on, Editors-in-Chief are listed with the year they took office. They continued to serve until the next date. (For a more detailed account see this *Monthly* 21, 1; 38, 305–320; 53, 582; Isis, 3, 490–491.)

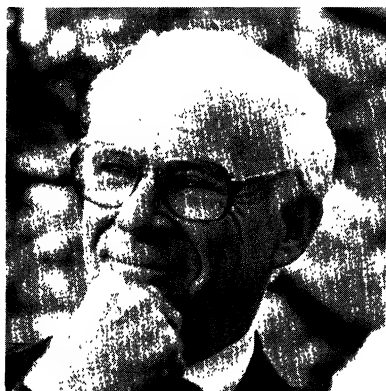
Early Boards of Editors: 1894–1915

1894	B. F. Finkel, J. M. Colaw
1903	B. F. Finkel, Leonard E. Dickson
1904	B. F. Finkel, Leonard E. Dickson, Saul Epstein
1905	B. F. Finkel, Leonard E. Dickson, Oliver E. Glenn
1907	B. F. Finkel, H. E. Slaught, Leonard E. Dickson
1909	B. F. Finkel, H. E. Slaught, G. A. Miller
1913	H. E. Slaught, E. R. Hedrick, G. A. Miller

Editors-in-Chief: 1916–1993

1916	Herbert Ellsworth Slaught <i>Univ of Chicago</i>	1952	Carl B. Allendoerfer <i>Univ of Washington</i>
1918	R. D. Carmichael <i>Univ of Illinois</i>	1957	Ralph D. James <i>Univ of B. C.</i>
1919	R. C. Archibald <i>Brown</i>	1962	Frederick A. Ficken <i>NYU</i>
1922	A. A. Bennett <i>Univ of Texas</i>	1967	R. A. Rosenbaum <i>Wesleyan</i>
1923	Walter Burton Ford <i>Michigan</i>	1969	Harley Flanders <i>Tel Aviv/Purdue</i>
1927	William Henry Bussey <i>Univ of Minnesota</i>	1974	Alex Rosenberg <i>Cornell</i>
1932	Walter B. Carver <i>Cornell</i>	1977	Ralph Boas <i>Northwestern</i>
1937	Elton J. Moulton <i>Northwestern</i>	1982	Paul R. Halmos <i>Indiana/Santa Clara</i>
1942	Lester R. Ford <i>Illinois Inst of Tech.</i>	1987	Herbert S. Wilf <i>U of Pennsylvania</i>
1947	Carroll V. Newsom <i>Oberlin</i>	1992	John H. Ewing <i>Indiana</i>

Recent Past Editors of The Monthly



R. A. Rosenbaum (1967–68)



Harley Flanders (1969–73)



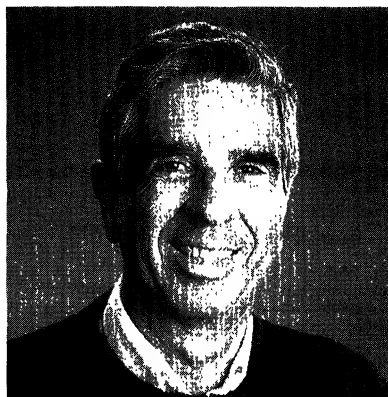
Alex Rosenberg (1974–77)



Ralph Boas (1977–1981)



Paul R. Halmos (1982–86)



Herbert S. Wilf (1987–91)

Quotients of Primes

David Hobby and D. M. Silberger

This note is motivated by our answer, appearing herein as Theorem 4, to the following question posed in a communication from our colleague William Vince Grounds: Given positive real numbers $a < b$, must there exist a pair p and q of prime integers such that $a < p/q < b$?

Henceforth \mathbb{R} denotes the set of all real numbers, \mathbb{R}^+ denotes the set of all positive real numbers, \mathbb{N} denotes the set of all positive integers, and \mathbb{P} denotes the set of all prime integers. When $S \subseteq \mathbb{N}$, then $\mathbf{F}(S)$ denote the set of all quotients p/q for which $\{p, q\} \subseteq S$ and $p \neq q$.

When a and b are real numbers then (a, b) denotes the *open interval* $\{t: a < t < b\}$, and $[a, b]$ denotes $\{t: a \leq t \leq b\}$. When $X \subseteq Y \subseteq \mathbb{R}$, then the set X is said to be *dense in Y under the usual topology on \mathbb{R}* if and only if $(a, b) \cap Y \neq \emptyset$ implies $(a, b) \cap X \neq \emptyset$ for every nonempty open interval (a, b) .

This notion of density occurs ubiquitously in point-set topology and analysis, and it has applications throughout mathematics; see [1, 2, 4]. For example, when D is a dense subspace of a topological space U , and when $f: D \rightarrow V$ is a continuous function, then f has a unique continuous extension $g: U \rightarrow V$.

When X is a topological space then the least cardinality of a dense subset of X is one among several interlocking measures of the ‘poverty’ of the topology on X . For instance, a topological space with a countable dense subset is called *separable*. These traits are ‘topological invariants’, and are fundamental to the classification of topological spaces.

It is well-known that \mathbb{R} under the usual topology is separable, and indeed that the denumerable set $\mathbf{F}(\mathbb{N})$ is dense in the subspace \mathbb{R}^+ . Theorem 4 states that $\mathbf{F}(\mathbb{P})$ is dense in \mathbb{R}^+ . If S is a finite subset of \mathbb{N} then $\mathbf{F}(S)$ is finite, and therefore $\mathbf{F}(S)$ fails to be dense in X for every infinite subspace X of \mathbb{R} .

Open Problem One. Characterize the family of all $S \subseteq \mathbb{N}$ for which $\mathbf{F}(S)$ is dense in \mathbb{R}^+ .

For $\{a, b\} \subseteq \mathbb{N}$ the expression $\mathbf{D}(a, b)$ denotes the set of all primes which occur as terms in the arithmetic progression $\langle a + jb \rangle_{j=0}^\infty$. P. G. Lejeune Dirichlet proved that when a and b are coprime then $\mathbf{D}(a, b)$ is infinite; see §8.4 of [3].

Open Problem Two. Is $\mathbf{F}(\mathbf{D}(a, b))$ dense in \mathbb{R}^+ whenever a and b are coprime?

Our interest in these two problems arises from Theorem 1, from Corollary 2, and from Theorem 4 in the paragraphs which follow. Theorem 1 indicates that there is a subset B of \mathbb{N} which is infinite, but which is nevertheless ‘sparse’ in the spirit of the present article.

Theorem 1. *There is an infinite set B of primes for which $\mathbf{F}(B)$ is not dense in \mathbb{R}^+ .*

Proof: Define $\mathbf{T}(n) := \mathbb{N} \cap [n, 2n]$ for each $n \in \mathbb{N}$. Bertrand's Postulate insures that for each $n \in \mathbb{N}$ we may choose a prime $\mathbf{b}(n)$ in the set $\mathbf{T}(2^{2n+1})$; see [3]. Let B denote $\{\mathbf{b}(n): n \in \mathbb{N}\}$.

Of course B is an infinite set of primes. So it suffices to show that $\mathbf{F}(B)$ is not dense in \mathbb{R}^+ . Now note that if $0 < i < j$ then $\mathbf{b}(i) < 2^{2i+2} < 2^{2j+1} < \mathbf{b}(j)$ whence $\mathbf{b}(i)/\mathbf{b}(j) < 2^{2i+2}/2^{2j+1} = 2^{2(i-j)+1} \leq 1/2$ while obviously $\mathbf{b}(j)/\mathbf{b}(i) > 2$. Therefore we have that $\mathbf{F}(B) \cap [1/2, 2] = \emptyset$. So $\mathbf{F}(B)$ is not dense in \mathbb{R}^+ . ■

The method used in the proof of Theorem 1 can be manipulated in order to secure the following stronger result:

Corollary 2. *There is an infinite set C of primes such that $\mathbf{F}(C)$ fails to be dense in each nonempty open interval $(a, b) \subseteq \mathbb{R}^+$.*

Proof: Let $0 < a < b$. For each $n \in \mathbb{N}$ let $\mathbf{c}(n)$ be a prime in the set $\mathbf{T}(2^{\mathcal{A}(n)})$ where $\mathcal{A}(n) := n^2$, and let $C := \{\mathbf{c}(n): n \in \mathbb{N}\}$. For each $n \in \mathbb{N}$ clearly $\mathbf{c}(n+1)/\mathbf{c}(n) > 2^{\mathcal{A}(n+1)}/2^{\mathcal{A}(n)+1} = 2^{2n}$, and also that $\mathbf{c}(n)/\mathbf{c}(n+1) < 2^{-2n}$. It follows that $\mathbf{F}(C) \cap (2^{-k}, 2^k)$ is a subset of the finite set $\mathbf{F}(\{\mathbf{c}(n): n \leq (k+1)/2\})$ for each $k \in \mathbb{N}$. So, since $(a, b) \subseteq (2^{-k}, 2^k)$ for some $k \in \mathbb{N}$, the set $\mathbf{F}(C) \cap (a, b)$ is at most finite. Therefore $\mathbf{F}(C)$ is not dense in (a, b) . ■

It is evident that the set $\mathbf{F}(C)$ of Corollary 2 is discrete, and that 0 is the only accumulation point of $\mathbf{F}(C)$.

Open Problem Three. For what sets $L \subseteq \mathbb{R}$ does there exist a set $\mathbf{A}(L) \subseteq \mathbb{N}$ such that L is the set of all accumulation points of the set $\mathbf{F}(\mathbf{A}(L))$? For what sets $M \subseteq \mathbb{R}$ does there exist a set $\mathbf{E}(M) \subseteq \mathbb{N}$ such that M is the closure of the set $\mathbf{F}(\mathbf{E}(M))$? (See [2].)

Lemma 3. *Let $1 < \alpha \in \mathbb{R}$. Then there exists a positive integer $m(\alpha)$ such that $[\alpha^n, \alpha^{n+1}] \cap \mathbb{P} \neq \emptyset$ for every integer $n > m(\alpha)$.*

Proof: For $x \geq 2$ define $G(x) := \pi(x)\ln(x)/x$ where $\pi(x)$ denotes the number of primes which do not exceed x , and define $L(x) := \log_\alpha(G(x))$. For each $r \in \mathbb{N}$ such that $[\alpha^r, \alpha^{r+1}] \cap \mathbb{P} = \emptyset$ we have that $G(\alpha^r) = \pi(\alpha^r)r\ln(\alpha)/\alpha^r$ and that $G(\alpha^{r+1}) = \pi(\alpha^r)(r+1)\ln(\alpha)/\alpha^{r+1}$, whence $G(\alpha^{r+1})/G(\alpha^r) = (r+1)/r\alpha$. It follows for each such r that $L(\alpha^{r+1}) - L(\alpha^r) = \log_\alpha(G(\alpha^{r+1})/G(\alpha^r)) = \varepsilon(r) - 1$ where $\varepsilon(r) := \log_\alpha((r+1)/r)$.

Since $\lim_{x \rightarrow \infty} G(x) = 1$ by the Prime Number Theorem (see [3]), we have that $\lim_{x \rightarrow \infty} L(x) = 0$. Hence there exists $h_0 \in \mathbb{N}$ such that $L(\alpha^{r+1}) - L(\alpha^r) > -1/2$ for every integer $r > h_0$. But there exists also $h_1 \in \mathbb{N}$ such that $\varepsilon(r) < 1/2$ for every integer $r > h_1$. Let $m(\alpha) := \max\{h_0, h_1\}$. Then for every integer $r > m(\alpha)$ such that $\pi(x)$ is constant on $[\alpha^r, \alpha^{r+1}]$ we have that $L(\alpha^{r+1}) - L(\alpha^r) = \varepsilon(r) - 1 < -1/2$. But if $r > m(\alpha)$ then $L(\alpha^{r+1}) - L(\alpha^r) > -1/2$. So there is no integer $r > m(\alpha)$ for which $\pi(x)$ is constant on $[\alpha^r, \alpha^{r+1}]$. That is, $[\alpha^n, \alpha^{n+1}] \cap \mathbb{P} \neq \emptyset$ for every integer $n > m(\alpha)$. ■

Theorem 4. $\mathbf{F}(\mathbb{P})$ is dense in \mathbb{R}^+ .

Proof: It suffices to show for arbitrary real numbers $c > \varepsilon > 0$ that $[c - \varepsilon, c + \varepsilon] \cap \mathbb{F}(\mathbb{P}) \neq \emptyset$. Thus it suffices to show that $\mathbb{P} \cap [(c - \varepsilon)q, (c + \varepsilon)q] \neq \emptyset$ for some prime q . So choose such c and ε , then choose any real number $\alpha > 1$ such that $\alpha^2 < (c + \varepsilon)/(c - \varepsilon)$, and finally choose a prime $q > \alpha^{m(\alpha)}/(c - \varepsilon)$ where $m(\alpha)$ is as in Lemma 3. We argue that $\log_\alpha((c + \varepsilon)q) - \log_\alpha((c - \varepsilon)q) = \log_\alpha((c + \varepsilon)/(c - \varepsilon)) > \log_\alpha(\alpha^2) = 2$. So $[n, n + 1] \subseteq [\log_\alpha((c - \varepsilon)q), \log_\alpha((c + \varepsilon)q)]$ for some integer n , whence $[\alpha^n, \alpha^{n+1}] \subseteq [(c - \varepsilon)q, (c + \varepsilon)q]$. Since $\alpha^{m(\alpha)} < (c - \varepsilon)q < \alpha^n$, we also have that $n > m(\alpha)$. So $[\alpha^n, \alpha^{n+1}] \cap \mathbb{P} \neq \emptyset$ by Lemma 3, whence $[(c - \varepsilon)q, (c + \varepsilon)q] \cap \mathbb{P} \neq \emptyset$ as required. ■

Notice that the proof of Theorem 4 guarantees that for every pair ε and c of positive real numbers with $\varepsilon < c$ there is a smallest prime $k(c, \varepsilon)$ such that for every prime $q \geq k(c, \varepsilon)$ there is a prime p for which $c - \varepsilon \leq p/q \leq c + \varepsilon$. What is the best easily characterized lower bound on the function $k(c, \varepsilon)$?

None of the results above depend in an obvious fashion upon the arithmetic properties of prime numbers, but depend instead upon their distribution in the sequence $1, 2, 3, \dots$. So, observations like ours are indubitably related to the number theoretic and probabilistic densities of a sequence of positive integers. These notions are discussed for instance in Chapter 11 of [3].

ACKNOWLEDGMENT. We are indebted to the referee for suggestions leading to improvements in our exposition.

REFERENCES

1. E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, New York, 1965.
2. J. L. Kelley, *General Topology*, Van Nostrand, New York, 1955.
3. I. Niven, H. S. Zuckerman, and H. L. Montgomery, *An Introduction to the Theory of Numbers*, fifth edition, Wiley, New York, 1991.
4. G. F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York, 1963.

*Department of Mathematics
The State University of New York
New Paltz, NY 12561*

*Department of Mathematics
The State University of New York
New Paltz, NY 12561*

and

*Universidade Federal de Santa Catarina
88000-Florianópolis-Santa Catarina
Brasil*

Notice About Forwarding the MONTHLY

A member or subscriber wishing to have his copy of the June–July MONTHLY forwarded to a summer address should leave four cents postage with his local post office or carrier, with instructions for remailing.

—*American Mathematical Monthly*
45, (1938) p. 331.

Pathological Functions for Newton's Method

George C. Donovan, Arnold R. Miller*
and Timothy J. Moreland

In the solution of equations by numerical methods, a commonly used stopping criterion is

$$|x_{n+1} - x_n| < \varepsilon, \quad (1)$$

where x_n is the n th term of the sequence generated by the method, and $\varepsilon > 0$ is the tolerance. For a specific method, the bisection method, it is not difficult to show that criterion (1) can never fail: if (1) is satisfied, then we also have $|x_{n+1} - x^*| < \varepsilon$, where x^* is the root. However, as the widely used text in numerical analysis by Burden and Faires [1] points out, in general, criterion (1) can fail. The text's argument is based only on abstract considerations, namely, that there exist sequences (e.g., the partial sums of the harmonic series) for which (1) is true but which nonetheless diverge. An example of a function and numerical method generating a sequence having this property is not given.

In this paper, we derive two functions, which exhibit this "false convergence" phenomenon. The first of these has no real root, but nevertheless generates a sequence under Newton's method for which (1) is satisfied for any ε , namely, $\{\sqrt{u_n}\}$ where $u_n \in \mathbb{R}$, $u_{n+1} = u_n + 1$, and $n = 0, 1, \dots$. Although this sequence satisfies (1), it obviously does not converge. The second function, like the first one, appears to converge where there are no roots, but it has a real root, to which Newton's method will never converge.

DERIVATION. To generate the sequence $\{\sqrt{u_n}\}$ we require a function f such that (Newton's method)

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

where $x_n = \sqrt{u_n}$, $u_{n+1} = u_n + 1$, $u \in \mathbb{R}$, and $n = 0, 1, \dots$. Rearranging this equation and utilizing the fact that $u_{n+1} = u_n + 1$ gives the differential equation

$$f'(x_n) = \frac{df}{dx_n} = \frac{f(x_n)}{x_n - \sqrt{x_n^2 + 1}}$$

or

$$\frac{df}{f} = \frac{dx}{x - \sqrt{x^2 + 1}} = (-x - \sqrt{x^2 + 1}) dx, \quad (2)$$

*Author to whom inquiries should be sent.

which has a particular solution

$$f(x) = \frac{c \exp\left[-\frac{1}{2}(x^2 + x\sqrt{x^2 + 1})\right]}{\sqrt{x + \sqrt{x^2 + 1}}}, \quad (3)$$

where c is the constant of integration.

By inspection of (3), f is bounded above by ce^{-x^2} for large values of x .

We now derive a function h that has a zero at $x = 0$, but nonetheless exhibits the above pathological behavior. For h , Newton's method should generate a divergent sequence $\{x_n\}$ for every starting value other than $x_0 = 0$, but should eventually satisfy criterion (1) for sufficiently large n . Thus, for x near zero, we chose for h to behave like the function $\sqrt[3]{x}$ in that Newton's method has a repelling fixed point at zero. For large x , we chose for h to behave like the function e^{-x^2} so that the sequence generated by Newton's method will resemble the one generated for function f , and will exhibit the "false convergence" property. We therefore make h the product,

$$h(x) = \sqrt[3]{x} e^{-x^2},$$

which, as we will demonstrate, has the desired properties. Note that because of the way that h is defined, it is bounded below by e^{-x^2} .

PROPERTIES OF FUNCTIONS

Function f . Function f , defined by equation (3) and generating the sequence $\{\sqrt{u_n}\}$, is strictly decreasing, since the first derivative of f is

$$f'(x) = f(x)(-x - \sqrt{x^2 + 1}) < 0. \quad (4)$$

In the limits of $x \rightarrow -\infty$ and $x \rightarrow +\infty$, we have $f(x) \rightarrow +\infty$ and $f(x) \rightarrow 0$, respectively. To demonstrate the limit as $x \rightarrow -\infty$, consider first the limit of the exponent in (3):

$$\begin{aligned} \lim_{x \rightarrow -\infty} -\frac{1}{2}(x^2 + x\sqrt{x^2 + 1}) \\ &= \lim_{x \rightarrow -\infty} -\frac{1}{2} \left[(x^2 + x\sqrt{x^2 + 1}) \cdot \frac{x^2 - x\sqrt{x^2 + 1}}{x^2 - x\sqrt{x^2 + 1}} \right] \\ &= \lim_{x \rightarrow -\infty} -\frac{1}{2} \left(\frac{x^4 - x^2(x^2 + 1)}{x^2 - x\sqrt{x^2 + 1}} \right) = \lim_{x \rightarrow -\infty} \frac{1}{2} \left(\frac{x^2}{x^2 - \sqrt{x^4 + x^2}} \right) = \frac{1}{4}. \end{aligned}$$

Therefore, in the limit as $x \rightarrow -\infty$, the numerator of (3) is $ce^{1/4} > 0$. In the denominator, we have

$$\lim_{x \rightarrow -\infty} x + \sqrt{x^2 + 1} = \lim_{x \rightarrow -\infty} \frac{-1}{x - \sqrt{x^2 + 1}} = 0$$

where the limit approaches 0 from above. Hence, $\lim_{x \rightarrow -\infty} f(x) = +\infty$. The graph of f is shown in Fig. 1.

Function h . The function $h(x) = \sqrt[3]{x} e^{-x^2}$ has a single zero at $x = 0$, but as we will see, iteration under Newton's method can never converge to this zero unless we are lucky enough to choose $x_0 = 0$. However, if the convergence criterion $|x_{n+1} - x_n| < \varepsilon$ is used (and if none of the iterates is a critical point of h , in which

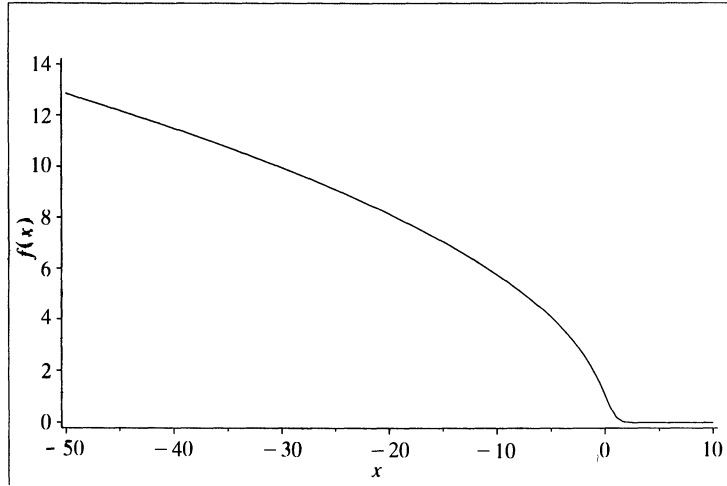


Figure 1. Function f .

case the sequence would diverge), Newton's method will appear to converge in one of the tails of the function, i.e., where there are no zeroes.

Applying Newton's method to h yields

$$x_{n+1} = x_n - \frac{h(x_n)}{h'(x_n)} = x_n - \frac{x_n}{\frac{1}{3} - 2x_n^2} = x_n \left[1 - \frac{1}{\frac{1}{3} - 2x_n^2} \right],$$

which is defined everywhere on the real line except at the critical points of h , which are $-1/\sqrt{6}$ and $1/\sqrt{6}$.

To show that the sequence $\{x_n\}$ cannot converge to zero, consider the case when x_n is in the interval $(-1/\sqrt{6}, 1/\sqrt{6})$ and $x_n \neq 0$. We see that

$$|x_{n+1}/x_n| = \left| 1 - \frac{1}{\frac{1}{3} - 2x_n^2} \right| > 2.$$

That is, instead of moving points closer to zero, this process drives them away, making each new point more than twice as far from zero as its predecessor. Clearly, this process eventually takes x_n outside the interval $(-1/\sqrt{6}, 1/\sqrt{6})$ for some n .

Now, suppose that x_k is outside of $(-1/\sqrt{6}, 1/\sqrt{6})$. We assume that x_k is in the right tail of the function, but the same properties hold for the left tail because the function is symmetric about the origin. Also, we assume that $x_k \neq 1/\sqrt{6}$, because the derivative of h at that point is zero, so Newton's method diverges.

Our convergence criterion is given as $|x_{n+1} - x_n| < \varepsilon$, which is the same thing as

$$\phi(x_n) = \frac{x_n}{2x_n^2 - \frac{1}{3}} < \varepsilon.$$

Clearly, $\phi(x_n)$ converges to zero as x_n goes to infinity, so there is some x , denoted x_c , such that for all $x_n \geq x_c$, the convergence criterion will be satisfied. To prove that x_n does get sufficiently large under iteration, we will assume that it does not and show that that leads to contradiction.

If $\phi(x_n)$ is never below ε (i.e., if x_n is never sufficiently large), then each iterate is greater than the last by at least ε . Thus, after $m > x_C/\varepsilon$ further iterations, we have

$$x_{k+m} \geq x_k + m\varepsilon > x_C.$$

This means that $\phi(x_{k+m}) < \varepsilon$, and contradicts the assumption that $\phi(x_n)$ was less than ε for no n . Thus, after sufficiently many iterations, x_n is large enough that $x_{n+1} - x_n < \varepsilon$, and the Newton's method algorithm stops, its criterion for convergence having been satisfied.

Unlike the function f , which is bounded above, $|h|$ is bounded below by e^{-x^2} for large values of x . A graph of h is shown in Fig. 2.

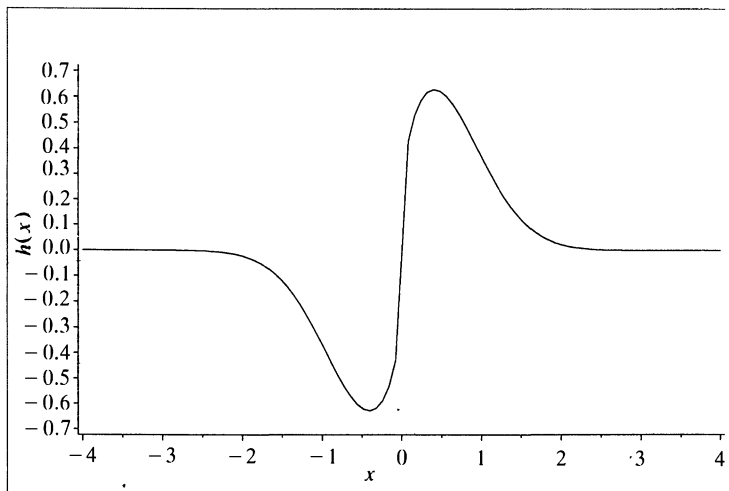


Figure 2. Function h .

THE GENERATED SEQUENCES. $\{\sqrt{u_n}\}$. Since function f in equation (3) is defined for all real numbers, any choice can be made for the initial approximation, x_0 , for Newton's method. Since function f was derived to generate the sequence $\{\sqrt{u_n}\}$, where $u_{n+1} = u_n + 1$, the numerical sequence is monotone and unbounded above. If the initial approximation $x_0 \geq 0$, then letting $x_0 = \sqrt{u_0}$, we get the sequence $\{\sqrt{u_n}\}_{n=0}^{\infty}$. If $x_0 < 0$, then from the iteration formula,

$$\begin{aligned} x_{n+1} &= x_n - \frac{1}{-x_n - \sqrt{x_n^2 + 1}} \\ &= x_n - \left(x_n - \sqrt{x_n^2 + 1} \right) \\ &= \sqrt{x_n^2 + 1}. \end{aligned}$$

It is interesting that, because f is concave for $x_0 < 0$, the second term of the sequence is independent of the sign of the first term, x_0 . Then, letting $x_1 = \sqrt{u_1}$, the remaining sequence is $\{\sqrt{u_n}\}_{n=1}^{\infty}$.

The "convergence" of the sequence is rather slow. If we used as a stopping criterion that the difference between succeeding terms be less than ε , i.e., $\sqrt{u_{n+1}} - \sqrt{u_n} < \varepsilon$, then $u_n \approx (1 - \varepsilon^2)^2/4\varepsilon^2$. As an illustration, if we let $x_0 = 1$

and $\varepsilon = 10^{-k}$, $k = 1, 2, 3, \dots$, then, neglecting round off error, the number of iterations to “convergence” would be $25 \times 10^{2(k-1)}$. In other words, to get, say, 3 decimal places of accuracy ($k = 3$) would require 250,000 iterations.

The sequence generated by h . The tails of h are very similar to the right tail of f , so the convergence of the sequence generated by h should resemble the convergence of $\{\sqrt{u_n}\}$.

Recall that for Newton’s method on h , the difference between consecutive iterates is

$$\phi(x_k) = \frac{x_k}{2x_k^2 - \frac{1}{3}}.$$

This expression is unwieldy, so for simplicity, we can obtain an estimate of ϕ by neglecting the $\frac{1}{3}$. Since x_k is large for large values of k , the constant $\frac{1}{3}$ is not significant. Thus, we have,

$$\phi(x_k) \approx \frac{x_k}{2x_k^2} = \frac{1}{2x_k}. \quad (5)$$

Using this estimate, we see that $\phi(x_k) < \varepsilon$ roughly when $x_k > 1/2\varepsilon$, so we have “convergence” for any such value of x_k .

Now, if we choose a typical starting value in the right tail, say $x_0 = 1$, we can find approximate upper and lower bounds on the number of iterations required to satisfy (1) for any value of ε . By (5), each iteration step-size for x_n between 1 and 2 is at least $1/4$, so it takes at most 4 steps to get from 1 to 2. Likewise, it takes at most 6 steps to get from 2 to 3, etc. In general, the number of iterations to get from 1 to m is at most

$$\sum_{i=2}^m 2i = m^2 + m - 2. \quad (6)$$

Similarly, it takes at least 2 steps to get from 1 to 2, 4 steps to get from 2 to 3, etc. Thus, the number of iterations required to get from 1 to m is at least

$$\sum_{i=1}^{m-1} 2i = m^2 - m. \quad (7)$$

Substituting $1/2\varepsilon$ for m in (6) and (7), we get

$$\frac{1}{4\varepsilon^2} - \frac{1}{2\varepsilon} \leq n \leq \frac{1}{4\varepsilon^2} + \frac{1}{2\varepsilon} - 2,$$

where n is the number of iterations to “convergence.”

The convergence for h is very similar to that for f . For example, to get 3 decimal places of accuracy ($\varepsilon = 10^{-3}$) would require approximately 250,000 iterations, which is what would be required for f .

As we saw above, e^{-x^2} is a bound (lower or upper) for the functions f and h for large values of x . Moreover, the three are related by the sequences generated by Newton’s method. Indeed, we have already seen that the iteration function for h is

$$x_{n+1} = x_n + \frac{x_n}{2x_n^2 - \frac{1}{3}},$$

and it is easy to check that the iteration function for e^{-x^2} is

$$x_{n+1} = x_n + \frac{1}{2x_n} = x_n - \frac{x_n}{2x_n^2 + 0}.$$

The function f , which is given in equation (3), is approximately

$$f(x) \approx \frac{ce^{-x^2}}{\sqrt{2x}},$$

for large values of x . Using this approximation yields the iteration function:

$$x_{n+1} \approx x_n + \frac{x_n}{2x_n^2 + \frac{1}{2}}.$$

These iteration functions differ from each other only by a constant in the denominator, which becomes insignificant for large values of x_n . So under iteration by Newton's method, f , h , and e^{-x^2} all behave similarly in the tails of the functions.

CONCLUSION. Functions f and h are useful as concrete examples demonstrating that stopping criterion (1) for the solution of equations can fail. Although neither of these functions ever converges to a root, they "converge" by this criterion. Indeed, it follows that they converge by the criterion

$$\frac{|x_{n+1} - x_n|}{|x_{n+1}|} < \varepsilon, \quad (8)$$

and, because f has the limit zero as $x \rightarrow +\infty$, and h has the limit zero as $x \rightarrow \pm\infty$, they also converge by the criterion

$$|\psi(x)| < \varepsilon,$$

where ψ is f or h . Hence, the functions converge by all standard stopping criteria. However, for reasonable values of ε , convergence is very slow. Besides this pathological behavior, the functions and their generated sequences exhibit other interesting behavior: the right tails of f and h resemble the right tail of the standard normal distribution. The second term, x_1 , of the sequence generated by f is independent of the sign of x_0 .

ACKNOWLEDGMENT. This work was subsidiary to and supported in part by a project funded by COPIC, Denver, Colorado. We thank Professor William Dorn, University of Denver, and Professor William Briggs, University of Colorado at Denver, for reading the manuscript and making suggestions for improvement. We also thank a referee and the journal editor for making suggestions for improvement.

REFERENCE

1. R. L. Burden and J. D. Faires, *Numerical Analysis*, 4th ed., PWS-Kent, Boston, 1989.

*Department of Mathematics and Computer Science
University of Denver
Denver, CO 80208*

NOTES

Edited by: **John Duncan**

THE NOTES SECTION IS BACK!

By popular demand (that charming euphemism for creative decisions by the editor-in-chief) the *Monthly* has resurrected the section for Notes. The inside cover continues to carry a succinct definition of what constitutes a typical Note, but this is an appropriate juncture for a brief essay on the subject.

A Note is short in length but high in content. It may be a clever new proof of an old theorem that adds insight as well as cutting out the use of heavy mathematical machinery; referees often comment that they will incorporate such proofs into their future teaching. Thoughtful instructors may thus incorporate into their course a theorem that was earlier considered too hard or advanced (an example from the past is Zagier's one sentence proof of Fermat's two squares theorem). On the other hand, when a theorem already has a transparent and simple proof, we are unlikely to accept an arcane proof even though it appealed greatly to the author.

A Note may be a new gem (say, the resolution of a conjecture about the triangle). It may be a solution of a problem that was posed in a recent *Monthly* article. It may be an unexpected twist to an old problem, perhaps adding surprising quantification to a qualitative result.

Of course we consider Notes on calculus (many thousands of our readers teach the subject year after year and appreciate good new ideas), but a high percentage of submitted Notes on calculus are not suitable. Often a proof that purports to bypass the completeness axiom already uses an intuitive idea that needs the completeness axiom for its justification. Sometimes an idea is well known but has just been discovered by an author who is anxious to tell others. Decisions on such articles can be difficult.

A Note might evoke one of the following reactions. "That's delightful". "I wish I had worked on that problem; I would have found that solution." "I could never have come up with that idea." This last comment needs to be laid alongside a similar comment made by mathematicians who re-read something they did many years ago. "How did I ever manage to do that?" The answer is—by wrestling with the problem... for a long time. We certainly hope that the Notes Section will encourage readers to keep on wrestling with half-completed projects.

Notes are refereed. Occasionally the editor will make a preliminary intervention to coax a good idea that needs more work. At present we receive a large number of manuscripts. Although our rejection rate is high we try to redirect some articles that would fit better as short notes in a research journal. We try to achieve a balance of themes in each issue. In the course of this first year, the nature of acceptable Notes ought to become clear to our readers. The flavor will reflect in part the personality of the Notes Editor, who takes the ultimate responsibility—and plaudits and brickbats.

John Duncan, Notes Editor

A Short Proof of a Theorem of Erdős and Mordell

André Avez

In 1935, Paul Erdős conjectured that for any point I inside (or on the boundary of) a triangle ABC , the sum of the distances from I to the vertices is at least twice the sum of the distances from I to the sides of $\triangle ABC$. He further conjectured that equality would hold if and only if $\triangle ABC$ is equilateral and I is its circumcenter.

Though this is easy to state and understand, the first proof was discovered only in 1937, by L. J. Mordell. It is by no means an elementary one. The first elementary proof was found by D. K. Kazarinoff in 1945 (see his son's book [2]). It is so tricky that it seems artificial.

The purpose of this note is to give a proof which seems natural and is accessible to college students.

We need two preliminary results. First is the elementary fact that $r + r^{-1} \geq 2$ for every $r > 0$, with equality if and only if $r = 1$ (to see this, expand the left side of $(r - 1)^2 \cdot r^{-1} \geq 0$). The second result is known as Ptolemy's theorem: *Let $ABCD$ be a convex quadrilateral inscribed in a circle. Then the sum of the products of the opposite sides is equal to the product of the diagonals:*

$$AC \cdot BD = AB \cdot CD + BC \cdot DA.$$

One can give an elementary proof of this via inversion. Another pleasant one, using complex numbers, can be found in [1], p. 105.

Proof of the Erdős-Mordell theorem. Let I be a point inside (or on the boundary of) $\triangle ABC$. Let the distances from I to the vertices be $a = IA$, $b = IB$ and $c = IC$, and the distances from I to sides BC , CA , and AB be u , v and w respectively. Let S be the circle passing through the three vertices A , B and C . Suppose the line through A and I also meets the circle S at A' . Ptolemy's theorem applied to $ABA'C$ gives

$$A'C \cdot AB + BA' \cdot AC = AA' \cdot BC. \quad (1)$$

Now let IH be the altitude of $\triangle AIC$ and let A'' be the point which is diametrically opposite A' on S . Since the inscribed angles $\angle A'AC$ and $\angle A'A''C$ are equal, the right triangles $\triangle AIH$ and $\triangle A''A'C$ are similar. Therefore

$AI \cdot A'C = A'A'' \cdot IH = v$ if, for convenience, we take the diameter of S equal to 1 (since then $A'A'' = 1$ and $IH = v$ in any case). By similar reasoning we can conclude that $IA \cdot BA' = w$. Multiplying both sides of (1) by $IA = a$ and dividing by BC , we get

$$v \frac{AB}{BC} + w \frac{AC}{BC} = aAA'.$$

We can apply this same reasoning to the other two vertices to get two more corresponding equalities. Adding the right and left sides, respectively of these

equalities yields

$$a \cdot AA' + b \cdot BB' + c \cdot CC' = u \left(\frac{AC}{AB} + \frac{AB}{AC} \right) + v \left(\frac{BA}{BC} + \frac{BC}{BA} \right) + w \left(\frac{CA}{CB} + \frac{CB}{CA} \right).$$

The left-hand side is less than or equal to $a + b + c$, with equality holding if and only if $AA' = BB' = CC' = 1$, which means that AA' , BB' and CC' are diameters of S , i.e., that I is the circumcenter of S . On the other hand, our first preliminary result implies that the right side is greater than or equal to $2(u + v + w)$, with equality holding if and only if $AB = AC = BC$. The theorem is proved.

Corollary.

$$2 \cdot \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \right) \leq \frac{1}{u} + \frac{1}{v} + \frac{1}{w}.$$

Proof: Transform the figure by polar reciprocity with respect to the circle of center I and radius 1.

Complex number interpretation. Let I be the origin in the complex plane \mathbf{C} and identify any point of \mathbf{C} with its complex number representation.

If I , A and B are distinct, then by minimizing $|tA + (1 - t)B|$ for $t > 0$, we obtain the distance from I to the line AB :

$$w = \frac{|\overline{A}B - A\overline{B}|}{2|A - B|}$$

and similar expressions for u and v . Now the Erdős-Mordell theorem reads:

$$\frac{|\overline{A}B - A\overline{B}|}{|A - B|} + \frac{|\overline{B}C - B\overline{C}|}{|B - C|} + \frac{|\overline{C}A - C\overline{A}|}{|C - A|} \leq |A| + |B| + |C|,$$

with equality holding if and only if A , B and C are the three roots of the equation $z^3 - A^3 = 0$. I know of no direct proof of this result. As a corollary, by changing A , B and C into their inverses, we obtain a new form of the Erdős-Mordell theorem:

$$2(ua + vb + wc) \leq bc + ca + ab.$$

Higher dimensions. Equality occurs in the Erdős-Mordell theorem when the configuration A, B, C, I possesses the highest degree of symmetry. What about dimension 3? Considering the regular tetrahedron suggests the following: given any point I inside (or on the boundary of) a tetrahedron $ABCD$, the sum S of the distances from I to the vertices is at least three times the sum s of the distances from I to the faces. But this is not true. For a counterexample, take $AC = AD = BC = BD$, $\angle ACB = \angle BDA = \text{right angle}$, C close to D and I the midpoint of AB . Then $S = IA + IB + IC + ID = 2 \cdot AB$, and s is close to $AB/\sqrt{2}$. Therefore S/s is close to $2\sqrt{2}$ which is less than 3. This is an interesting breaking of symmetry.

ACKNOWLEDGMENT. This research was supported by the School of Mathematics, University of Minnesota, which made possible my visit through an Ordway Endowment. For the translation into English I enjoyed the help of Prof. Leon Green.

REFERENCES

1. G. H. Hardy, *A Course of Pure Mathematics* (9th edition), Cambridge University Press, 1948.
2. N. Kazarinoff, *Geometric inequalities*, Random House, 1961.

*Department of Mathematics
University of Paris VI
2 Place Jussieu, 75005 Paris
France*

The Computer Solves the Three Tower Problem

Arthur Engel

Consider the following probability problem: *We have three piles with a, b, c chips, respectively. Each second a pile X is selected at random, then another pile Y is chosen at random and a chip is moved from X to Y . Find the expected waiting time $f(a, b, c)$ until one pile is empty. (Three Tower Problem, or TTP.)*

This problem is due to Lennart Råde from Gothenburg University, Sweden. During the last 20 years he posed it to numerous people, but nobody could solve it [5]. I heard of it during a Statistics Conference in New Zealand 1990. It became a simulation exercise in a book I was writing at the time [3]. The simulation problem gives numerical answers to specific inputs a, b, c . On January 17, 1992 I loaded the program again and started to experiment. In 15 minutes I guessed the formula

$$f(a, b, c) = \frac{3abc}{a + b + c}. \quad (1)$$

Once you have guessed the formula, the proof is a routine matter. Start in state (a, b, c) . In one step you are in one of the neighboring states $(a, b + 1, c - 1)$, $(a, b - 1, c + 1)$, $(a + 1, b, c - 1)$, $(a - 1, b, c + 1)$, $(a + 1, b - 1, c)$, $(a - 1, b + 1, c)$ with the same probability $1/6$. So we have

$$f(a, b, c) = 1 + \frac{1}{6} \sum f(x, y, z) \quad (2)$$

over all neighbors (x, y, z) of (a, b, c) with boundary conditions

$$f(a, b, 0) = f(a, 0, c) = f(0, b, c) = 0. \quad (3)$$

It looks pretty hopeless to solve the functional equation (2) with the boundary conditions (3). But thanks to the PC we have the guess (1). It obviously satisfies (3) and a short calculation shows that (2) is also satisfied. So we have a solution to our problem. Its uniqueness can be proved by a standard argument, which we reproduce to make the paper self contained. See [1] or [2].

Suppose that $g(a, b, c)$ is another solution. Consider $h(a, b, c) = f(a, b, c) - g(a, b, c)$. Then

$$h(a, b, c) = \frac{1}{6} \sum h(x, y, z) \quad (4)$$

over all neighbors of (a, b, c) .

The function h is defined for finitely many points. At some of these points h assumes its maximum M . Because of (4) $h(x, y, z) = M$ for all six neighbors of (a, b, c) . And their neighbors have also the same h -value M , and so on, until we reach the boundary, at which h has value 0. Thus $h(a, b, c) \leq 0$ everywhere. Similarly we can show that $-h \leq 0$. Thus $h = 0$, and $f(a, b, c) = g(a, b, c)$ everywhere. So f is unique.

Lecturing in Norway, Råde mentioned that the TTP has recently been solved by me. A listener asked about the expected duration $g(a, b, c)$ of the following modification of the TTP: *Start with three towers. As soon as one tower is empty continue playing with two players until just one is left.* At the lecture it was agreed that this would be a harder problem to solve. Råde challenged me to find $g(a, b, c)$. He added that he also would like to know the probability p_a that the a -tower first becomes empty.

With my PC I started to work empirically on $g(a, b, c)$. Instead of 15 minutes it took me several hours of hard work. The trouble was that I was looking for a more complicated formula. At the end I found the much simpler correct formula

$$g(a, b, c) = ab + bc + ca. \quad (5)$$

It is easy to show that (2) is satisfied when $a, b, c > 0$. The new boundary conditions are

$$g(a, b, 0) = ab, \quad g(a, 0, c) = ac, \quad g(0, b, c) = bc. \quad (6)$$

(6) is the expected duration for the Two Tower Problem. This is a classic result, which is equivalent to the Gamblers Ruin Problem. See [4]. Had I looked at (6) first, they would have immediately suggested (5).

By analogy I was able to write down the solution of the modified n -Tower-Problem:

$$g(x_1, \dots, x_n) = \sum_{i < k} x_i x_k. \quad (7)$$

It was also easy to guess the following version of Råde's second problem: The i th tower is the winner (in the game which continues until one pile is left) with probability

$$p_i = \frac{x_i}{x_1 + \dots + x_n}. \quad (8)$$

Both formulas (7) and (8) satisfy the appropriate recurrences and boundary conditions. The proof of (7) involves induction on n . The boundary conditions for a given value of n are determined by the solution of the problem for $n - 1$. Uniqueness is proved as in the case of the Three-Tower-Problem.

A related result could also be found with my PC searching for several hours: *Players 1, 2, 3 start with a, b, c chips, respectively. In one round each player stakes one chip. Then a 3-sided symmetric die labeled 1, 2, 3 is rolled and the winner gets all the chips staked. If a player is broke the game continues with two players until one player has accumulated all the chips.* The expected number of rounds is

$$h(a, b, c) = ab + bc + ca - \frac{2abc}{a + b + c - 2}. \quad (9)$$

If the game stops as soon as one tower is empty the expected duration is

$$h(a, b, c) = \frac{abc}{a + b + c - 2}. \quad (10)$$

This result was communicated to me by a former IMO contestant Michael Stoll. It was found ten years ago during a summer academy for gifted high school students. Despite huge efforts they were unable to handle four players.

The original Four Tower Problem is still unsolved. I experimented extensively for many hours, but all my guesses turned out to be wrong. $f(a, b, c, d)$ seems to be a very complicated function, as can be seen from the exact value $f(3, 2, 2, 2) = 350612/69969$. No simple formula can give such a complicated result for so small values of a, b, c, d . The only thing I could do was to guess a good approximation

$$f(a, b, c, d) \approx \frac{6abcd}{ab + ac + ad + bc + bd + cd}. \quad (11)$$

It is easy to see that $f(a, b, c, d)$ has the form $p(a, b, c, d)/q(a, b, c, d)$ with polynomials which are symmetric in a, b, c, d . In addition q seems to be constant, depending only on $a + b + c + d$. The use of *Mathematica* may bring more success. I worked numerically with Turbo Pascal as in [3].

REFERENCES

1. P. G. Doyle and J. L. Snell, *Random Walks and Electrical Networks*, Carus Monograph #22. Math. Assoc. America 1984, pp. 17, 18.
2. E. B. Dynkin and A. A. Yushkevich, *Markov Processes; Theorems and Problems*, Plenum Press, 1969, Ch. 1 (especially the exercises 18–24).
3. Arthur Engel, *Exploring Mathematics with Your Computer*, NML 35, MAA, 1992.
4. W. Feller, *An Introduction to Probability and Its Applications*, vol. 1, Sect. 14.3.
5. Lennart Råde, *Take a Chance with Your Calculator*, Dillithium Press, p. 17.

*Department of Mathematics
University of Frankfurt
Frankfurt / M.
Germany*

A Linear Algebra Approach to Cyclic Extensions in Galois Theory

Evan G. Houston

A beginning course in Galois theory often includes a discussion of cyclic extensions, that is, Galois extensions whose Galois groups are cyclic. The usual approach (see, e.g., [1] and [2]) is to derive the results on cyclic extensions as

corollaries to the “norm” (Hilbert’s “Theorem 90”) and “trace” theorems. In this note we offer an alternate approach based on linear algebra. Notation and ideas are standard as in [1] and [2]. All field extensions considered are finite.

Theorem 1 (cf. [2, Theorem 34]. *Let n be a positive integer, let K be a field of characteristic 0 or characteristic p with $(p, n) = 1$, and assume that K contains a primitive n th root of unity. If F/K is a Galois extension with cyclic Galois group of order n , then $F = K(\alpha)$, where α is a root of an irreducible polynomial $x^n - a \in K[x]$.*

Proof: We content ourselves with producing an element α in F for which $\sigma(\alpha) = \omega\alpha$, where ω is a primitive n th root of unity in K and σ generates the Galois group; it is then relatively easy to show that the element α does what is needed. The usual approach is to note that the norm of ω is 1 and then to invoke the norm theorem (Hilbert’s Theorem 90). Instead, view σ as a linear operator on the K -vector space F . The trick is to show that ω is an eigenvalue of σ . Since $\sigma^n = id$, σ satisfies the polynomial $x^n - 1$. By independence of characters ([2, Theorem 30] or [1, Lemma V.7.5]), σ does not satisfy a polynomial of degree less than n . It follows that $x^n - 1$ is the minimum, and therefore also the characteristic, polynomial of σ . Hence ω is indeed an eigenvalue. \square

Theorem 2 (cf. [2, Theorem 32] or [1, Proposition V.7.8]). *Let K be a field of characteristic p , and let F/K be a Galois extension with cyclic Galois group of order p . Then $F = K(\alpha)$, where α is a root of an irreducible polynomial $x^p - x - a \in K[x]$.*

Proof: The hard part is to produce an element $\alpha \in F$ with $\sigma(\alpha) - \alpha = 1$, where σ is a generator of the Galois group. This can be done without the trace theorem as follows. Again, view σ as a linear operator, and define another linear operator $\tau = \sigma - id$. We wish to show that $1 \in \text{im}(\tau)$. Since F/K is a Galois extension, $\ker(\tau) = K$. Claim: $\text{im}(\tau) \cap \ker(\tau) \neq \{0\}$. Otherwise, we have $F = \text{im}(\tau) + \ker(\tau)$. Now $\sigma^p = id$ implies that τ^p is the zero map. Let $\beta \in F$, and write $\beta = \gamma + \tau(\delta)$, where $\gamma \in \ker(\tau)$. Then $\tau^{p-1}(\beta) = \tau^{p-1}(\gamma) + \tau^p(\delta) = 0$, whence τ^{p-1} is the zero map. By induction, we get that τ is the zero map, which is nonsense. Hence the claim is true. Since $\ker(\tau) = K$ has dimension 1, it follows that $\text{im}(\tau) \cap \ker(\tau) = K$, and so $K \subseteq \text{im}(\tau)$. In particular, $1 \in \text{im}(\tau)$, as desired. \square

We close with a proof of the trace theorem which uses just independence of characters and the rank-plus-nullity theorem from linear algebra. Recall that for a Galois extension F/K the trace is defined as follows: for $\alpha \in F$, $T(\alpha) = \sum \sigma(\alpha)$, the sum being taken over all elements σ of the Galois group of F/K .

Theorem 3 (cf. [2, Theorem 31] or [1, Theorem V.7.6 (i)]). *Let F/K be a Galois extension with cyclic Galois group, generated by σ . Then for $\alpha \in F$, $T(\alpha) = 0$ if and only if $\alpha = \sigma(\beta) - \beta$ for some $\beta \in F$.*

Proof: Put $\tau = \sigma - id$. Then both T and τ are linear operators on the K -vector space F , and the proof boils down to showing that $\ker(T) = \text{im}(\tau)$. It is easy to see

that $\text{im}(\tau) \subseteq \ker(T)$. Since F/K is a Galois extension, we have that $\ker(\tau) = K$, so that $\text{im}(\tau)$ is $(n - 1)$ -dimensional, where $n = [F:K]$. By independence of characters, T is not the zero map, so that $\ker(T)$ also has dimension $n - 1$. It follows that $\ker(T) = \text{im}(\tau)$, completing the proof. \square

REFERENCES

1. T. W. Hungerford, *Algebra*, Springer-Verlag, New York, 1974.
2. I. Kaplansky, *Fields and Rings*, Chicago Lectures in Mathematics Series, University of Chicago Press, Chicago, 1972.

Department of Mathematics
University of North Carolina at Charlotte
Charlotte, NC 28223
fma00egh@unccvm.bitnet

UNSOLVED PROBLEMS

Edited by: Richard Guy

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

When Does a Polynomial over a Finite Field Permute the Elements of the Field?, II

Rudolf Lidl and Gary L. Mullen

In [16] we provided a brief survey of the main known classes of **permutation polynomials** (PPs) over finite fields, ($f \in F_q[x]$ is a PP if f induces a 1-1 mapping on F_q where F_q is the finite field of order q , with q a prime power). In addition, nine problems concerning PPs were discussed in [16] and as a result it seems that [16] has at least partly motivated a number of subsequent papers [2-5, 7, 9-11, 23-25] in which various aspects of PPs have been considered.

Progress has recently been made on a number of the problems raised in [16] and so it seems timely to provide a brief progress report surveying these developments. In addition we discuss an additional set of problems in the hope that these will likewise spur readers to develop further results concerning PPs.

1. Progress Report. We briefly describe some results that have recently appeared and begin by listing two major breakthroughs. We use the same numbering of the problems as in [16].

P8. Chowla and Zassenhaus conjecture. *If p is a sufficiently large prime and $f(x)$ of degree ≥ 2 permutes F_p , then $f(x) + ax$ with $0 < a < p$ is not a PP of F_p .*

Cohen [2] has affirmatively resolved this conjecture. In fact even better, he proves the following refinement: Let $f(x)$ be a polynomial with integer coefficients and degree $n \geq 2$. Then, for any prime $p > (n^2 - 3n + 4)^2$ for which f (considered modulo p) is a PP of degree n of F_p , there is no integer a with $1 \leq a < p$ for which $f(x) + ax$ is also a PP of F_p . This can be extended to tame PPs over general finite fields. Cohen's proof relies on the deep work of M. Fried in his proof [8] of Schur's conjecture which says that every integral polynomial which is a PP mod p

for infinitely many p is a composition of linear polynomials and Dickson polynomials. See also [22] for a recent survey of work on Schur's conjecture.

P9. Carlitz conjecture. *For each even positive integer k , there is a constant C_k such that for each finite field of odd order $q > C_k$, there does not exist a PP of degree k over F_q .*

By resolving singularities of plane curves over F_q , Wan [24] proved that the Carlitz conjecture is true for $k = 2r$ where r is an odd prime. Independently Cohen [3] has obtained the same result through the theory of primitive permutation groups. Cohen also proves the conjecture for each even $k < 1000$.

P2 of [16] asked for new classes of PPs. As a result of his study of factorable and exceptional polynomials over F_q , Cohen [4, 5] discovered the following new class of PPs. Let $L(x)$ be a linearized polynomial of the form $L(x) = \sum_{i=0}^k a_i x^{p^i}$ with the property that for some $s \geq 1$, $a_i = 0$ unless s divides i . Such an $L(x)$ is called a **p^s -polynomial**. Let d divide $p^s - 1$ where p does not divide d . Then $L(x) = xM(x^d)$ and $S(x) = xM^d(x)$ is called a **(p^s, d) -polynomial**. If M has no roots in F_q then S is a PP of F_q , see Cohen [4]. As an example, take $q = 2$ and define F_8 as $F_2(\alpha)$, where $\alpha^3 = \alpha + 1$. Then $M(x) = x^5 + (\alpha^2 + \alpha + 1)x + \alpha + 1$ in $F_8[x]$ factors as $M(x) = (x^2 + x + \alpha + 1)(x^3 + x^2 + \alpha x + 1)$ over F_8 into irreducible polynomials. Then $xM^3(x)$ is a PP over $F_{2^{3n}}$ for all integers n with $(n, 6) = 1$.

Wan & Lidl [25] studied another class of polynomials and showed for positive integers d and r satisfying $d|(q-1)$ and $f(x) \in F_q[x]$ that $x^r f(x^{(q-1)/d})$ is a PP of F_q if and only if the following conditions are satisfied: $(r, (q-1)/d) = 1$, $f(\omega^i) \neq 0$ for all $0 \leq i < d$ where $\omega = g^{(q-1)/d}$ denotes a primitive d th root of unity in F_q and g is a fixed primitive root of F_q and $\psi(f(\omega^i)/f(\omega^j)) \not\equiv r(j-i) \pmod{d}$, for all $0 \leq i < j < d$, where ψ is a multiplicative character with values in $\mathbb{Z}/d\mathbb{Z}$ such that for all $a \in F_q^*$, $\psi(a) \equiv \text{ind}_g(a) \pmod{d}$. Here $\text{ind}_g(a)$ is the residue class $b \pmod{d}$ such that $a = g^b$ and $\mathbb{Z}/d\mathbb{Z}$ denotes the ring of integers mod d .

In [9] von zur Gathen gave the following result concerning PPs and polynomials with large value sets. Let $f \in F_q[x]$ have degree $n \geq 1$, $|V_f|$ be the number of distinct images of f so that $|V_f| = |\{f(a) | a \in F_q\}|$, and let $\rho = q - |V_f|$. Then either $\rho = 0$ (so f is a PP) or $4n^4 > q$ or $2\rho n > q$. Hence if q is large and f is not a PP, then either n or ρ is large. See also [11].

P1 of [16] asked for a good algorithm to test whether a given polynomial is a PP of F_q . A probabilistic polynomial-time algorithm to test whether a given polynomial is a PP is given in von zur Gathen [10].

2. More Problems. We list further open problems and continue with the numbering of these unsolved problems from [16].

P10. For $p > 2$ a polynomial f is said to be **planar** if $f(x+e) - f(x)$ is a PP for every $e \in F_q^*$. As indicated in [6, 12, 21] such polynomials are important in the study of affine planes. Clearly any quadratic polynomial $ax^2 + bx + c$ is planar and it is shown in [12] and [21], that if $q > 2$ is prime, then every planar polynomial is quadratic. If $q = p^n$ with $n > 1$ there are planar polynomials over F_q of the form

$$\sum_{i,j=0}^{n-1} a_{ij} x^{p^i + p^j}, \quad (1)$$

where $a_{ij} \in F_q$. Prove or disprove the conjecture from [6] that every planar polynomial $f(x)$ with $f(0) = 0$ has the form (1).

P11. Let q be an odd prime, let $1 < k < q - 1$ with $k|(q - 1)$, and let $B = B(q, k)$ be the subgroup of F_q^* of order k . Assume that $f, g \in F_q[x]$ are both PPs and suppose that for each $x \in F_q$ and for each $b \in B$,

$$f(x + b) - g(x) \in B. \quad (2)$$

Such a pair (f, g) of PPs determines an automorphism of a design $D(q, k)$, see [15]. Conversely, each automorphism of $D(q, k)$ determines a pair (f, g) of PPs satisfying (2). Determine all pairs (f, g) of PPs satisfying (2).

P12. A PP $f(x)$ is called a **complete mapping** of F_q if $f(x) + x$ is also a PP. More generally, for a subset S of F_q containing 0, a polynomial with the property that $f(x) + ax$ is a PP for each $a \in S$ is called an **S-complete mapping**, see [1]. Complete mappings are useful in the study of orthogonal latin squares. The following conjecture has been proposed by Evans, Greene & Niederreiter [7]. If $f \in F_q[x]$ is such that $f(x) + ax$ is a PP for at least $\lfloor q/2 \rfloor$ values of $a \in F_q$, then $f(x) - f(0)$ is a linearized polynomial over F_q where $\lfloor \cdot \rfloor$ denotes the greatest integer function. By a linearized polynomial is meant a polynomial of the form $\sum_{i=0}^m a_i x^{p^i}$ where $q = p^n$. It is known from [7] that this conjecture is true for the case $q = p$ a prime and it is true for general q if $f(x) = x^e$. Prove or disprove the conjecture.

P13. Consider the binomial $f(x) = x^k + ax^j$ with $k > j \geq 1$, $\gcd(k, j) = 1$, and $a \in F_q^*$. It is shown in [23] that if $f(x)$ permutes F_q then $q \leq (k - 2)^4 + 4k - 4$ or $k = sp^r$ with $r \geq 1$ and q is a power of the characteristic p . For $k = 8$, $f(x)$ permutes F_q if and only if

- (i) $j = 1$ and $q = 2^{3r}$, $a^{(q-1)/7} \neq 1$ or $q = 29$, $a = \pm 4, \pm 10$,
- (ii) $j = 2$, $q = 2^{2r}$, $a^{(q-1)/3} \neq 1$,
- (iii) $j = 3$, $q = 11$, $a = \pm 2, \pm 4$,
- (iv) $j = 5$ and $q = 4$, $a \neq 1$ or $q = 7$, $a = \pm 3$.

Determine conditions on k, j and q so that $f(x)$ permutes F_q .

P14. If $k > j > i \geq 1$ with $\gcd(k, j, i) = 1$ and $a, b \in F_q^*$ with $a \neq b$, determine conditions on k, j, i and q so that $x^k + ax^j + bx^i$ permutes F_q .

P15. Characterize the PPs over F_q , $q = 2^e$, of the form $h_k(x) = 1 + x + x^2 + \cdots + x^k$. Matthews [18] showed that if $q = p^e$, with p odd, then $h_k(x)$ is a PP of F_q , if and only if $k \equiv 1 \pmod{p(q - 1)}$. When q is even, this condition is proved sufficient. Such PPs are useful in constructing ovals in projective planes $PG(2, q)$.

P16. Let $q > 2$ be even. Determine all PPs $f(x)$ over F_q with $f(0) = 0$ and $f(1) = 1$ such that for each $a \in F_q$ the polynomial f_a where $f_a(x) = (f(x + a) + f(a))/x$, $f_a(0) = 0$, is a PP. List all such PPs of degrees ≤ 6 similar to Dickson's list of all normalized PPs over F_q . See [13], [20] for a connection of these polynomials with hyperovals in $PG(2, q)$, see also [17, p. 504].

P17. In [14] it is shown that the existence of a j -plane is equivalent to the existence of a very special type of PP. In particular the authors give a construction which shows that if $x^2 + gx - f$ is irreducible over F_q , then there is an associated j -plane if and only if the polynomial $\phi_j(\mu, t) = ft(\mu(\mu + gt) - ft^2)^j$ is a PP for each $\mu \in F_q$. Several classes of j -planes are discussed in [14]. Find other classes. Such PPs are related to local PPs studied by Mullen [19] where $f(x, y)$ is local if $f(x, a)$ and $f(b, y)$ are PPs for all $a, b \in F_q$.

REFERENCES

1. W.-S. Chou, Permutation Polynomials over Finite Fields and Combinatorial Applications, Ph.D. thesis, Pennsylvania State University, 1990.
2. S. D. Cohen, Proof of a conjecture of Chowla and Zassenhaus on permutation polynomials, *Canad. Math. Bull.* 33 (1990), 230–234.
3. S. D. Cohen, Permutation polynomials and primitive permutation groups, *Archiv. Math.* (Basel), 57 (1991), 417–423 MRj:11145.
4. S. D. Cohen, Exceptional polynomials and the reducibility of substitution polynomials, *Enseignement Mathématique* 36 (1990), 53–65.
5. S. D. Cohen, The factorable core of polynomials over finite fields, *J. Austral. Math. Soc., Series A* 49 (1990), 309–318.
6. P. Dembowski & T. G. Ostrom, Planes of order n with collineation groups of order n^2 , *Math. Zeitschrift* 103 (1968), 239–258.
7. R. J. Evans, J. Greene & H. Niederreiter, Linearized polynomials and permutation polynomials of finite fields, *Michigan Math. J.*, to appear.
8. M. Fried, On a conjecture of Schur, *Michigan Math. J.* 17 (1970), 41–55.
9. J. von zur Gathen, Values of polynomials over finite fields, *Bull. Austral. Math. Soc.* 43 (1991), 141–146.
10. J. von zur Gathen, Tests for permutation polynomials, *SIAM J. Comput.* 20 (1991), 591–602.
11. J. von zur Gathen, Polynomials over finite fields with large images, ISSAC-90, Tokyo, Japan, *ACM Press* (1990), 140–144.
12. D. Gluck, A note on permutation polynomials and finite geometries, *Discrete Math.* 80 (1990), 97–100.
13. D. R. Glynn, A condition for the existence of ovals in $PG(2, q)$, q even. *Geom. Dedicata* 32 (1989), 247–252.
14. N. L. Johnson, R. Pomareda & F. W. Wilke, J -planes, *J. Combin Theory Set. A*, A56 (1991), 271–284.
15. W. M. Kantor, 2-Transitive designs, Combinatorics, *Proc. Adv. Study Inst. on Comb.*, Nijenrode Castle, Breukelen, Neth. (eds. M. Hall, Jr., and J. H. Van Lint) *Math. Centre Tracts* 57 (1974), 44–97, Math. Centrum, Amsterdam, 1974.
16. R. Lidl & G. L. Mullen, When does a polynomial over a finite field permute the elements of the field?, *Amer. Math. Monthly* 95 (1988), 243–246.
17. R. Lidl & H. Niederreiter, *Finite Fields*, Addison-Wesley, Reading, Massachusetts, 1983 (now distributed by Cambridge University Press).
18. R. W. Matthews, Permutation properties of the polynomials $1 + x + \cdots + x^k$ over a finite field, *Proc. Amer. Math. Soc.*, to appear.
19. G. L. Mullen, Local permutation polynomials over Z_p , *Fibonacci Quarterly*, 18 (1980), 104–108.
20. C. M. O’Keefe & T. Penttilä, A new hyperoval in $PG(2, 32)$. *J. Geometry* 44 (1992), 117–139.
21. L. Rónyai & T. Szőnyi, Planar functions over finite fields, *Combinatorica* 9 (1989), 315–320.
22. G. Turnwald, On Schur’s conjecture, *J. Austral Math. Soc., Series A*, to appear.
23. G. Turnwald, Permutation polynomials of binomial type, *Contributions to General Algebra* 6, Verlag Hölder-Pichler-Tempsky, Wien 1988, 281–286.
24. D. Wan, Permutation polynomials and resolution of singularities over finite fields, *Proc. Amer. Math. Soc.* 110 (1990), 303–309.
25. D. Wan & R. Lidl, Permutation polynomials of the form $x^r f(x^{(q-1)/d})$ and their group structure, *Monatsh. Math.* 112 (1991), 149–163.

Department of Mathematics
University of Tasmania, Hobart
Tasmania 7001
Australia
lidl@hilbert.maths.utas.edu.au

Department of Mathematics
Pennsylvania State University
University Park, PA 16802
mullen@math.psu.edu

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before June 30, 1993 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10274. *Proposed by Robert E. Byerly, Texas Tech University, Lubbock, TX.*

For odd integers n , let E_n be the smallest subfield of the real numbers closed under the function $x \mapsto \sqrt[n]{x}$.

(a) If $f(x) \in \mathbb{Z}[x]$ is irreducible in $\mathbb{Z}[x]$, show that $f(x)$ has at most one root in E_3 .

(b) Are there any such fields E_n and \mathbb{Z} -irreducible polynomials $f(x)$ for which $f(x)$ has more than one root in E_n ?

10275. *Proposed by Murray S. Klamkin and A. Liu, University of Alberta, Edmonton, Alberta, Canada.*

Let \mathcal{A} be a regular n -gon with edge length 2. Denote the consecutive vertices by A_0, \dots, A_{n-1} and introduce A_n as a synonym for A_0 . Let \mathcal{B} be a regular n -gon inscribed in \mathcal{A} with vertices B_0, \dots, B_{n-1} where B_i lies on $A_i A_{i+1}$ and $|A_i B_i| = \lambda < 1$ for $0 \leq i < n$. Also let C_i be the point on $A_i A_{i+1}$ with $|A_i C_i| = \alpha_i \leq \lambda$ for $0 \leq i < n$ and let \mathcal{C} denote the n -gon, also inscribed in \mathcal{A} , with vertices C_0, \dots, C_{n-1} .

With $P(\mathcal{F})$ denoting the perimeter of the figure \mathcal{F} , prove that $P(\mathcal{C}) \geq P(\mathcal{B})$.

10276. *Proposed by Stephen M. Gagola, Jr., Kent State University, Kent, OH.*

If

$$M = \begin{pmatrix} x & y \\ z & w \end{pmatrix},$$

define $\det M = xw - yz$ and $\text{dot } M = xz + yw$. Determine necessary and sufficient conditions on a field F , assumed to have characteristic different from 2, for the existence of quadratic forms $q_{ij} \in F[x, y, z, w]$ ($i, j \in \{0, 1\}$) such that $\det Q = (\det M)^2$ and $\text{dot } Q = (\text{dot } M)^2$, where

$$Q = \begin{pmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{pmatrix}.$$

In particular, do such forms exist when $F = \mathbb{Q}$?

10277. *Proposed by L. E. Mattics, University of South Alabama, Mobile, AL.*

Let p be a prime with $p \equiv 1 \pmod{4}$. Show that there are integers x and y such that $x^p + y^p$ is of the form $u^2 + pv^2$ for integers u and v , but $x + y$ is not of that form.

10278. *Proposed by Raphael M. Robinson, University of California, Berkeley, CA.*

A three-dimensional torus is formed from Euclidean 3-space by reducing each coordinate modulo 3. An *admissible* graph on this torus is one whose vertices are the 27 lattice points and whose edges are unit segments chosen so that exactly two meet at each vertex and form a right angle there. Find the total number of admissible graphs, taking account of position as well as of form.

10279. *Proposed by M. Al-Ahmar, Al-Fateh University, Tripoli, Libya.*

Let k and n be integers with $0 < k < n$, and let A be a real n by n orthogonal matrix with determinant 1. Let B be the upper left k by k submatrix of A , and let C be the lower right $(n - k)$ by $(n - k)$ submatrix of A .

- Show that $\det(B) = \det(C)$.
- Give a geometrical interpretation.
- Generalize to the case in which A is a unitary matrix.

10280. *Proposed by Donald E. Knuth, Stanford University, Stanford, CA.*

Define a random binary operation \star on the set $\{1, \dots, n\}$ by choosing every value independently, so that each of the n^{n^2} possible binary operations is equally likely.

- Prove that the axiom

$$((x \star x) \star x) \star ((x \star x) \star x) = x$$

holds for $1 \leq x \leq n$ with probability

$$\sum_{k=1}^n \frac{p_{n,k}}{n^{2n-k}},$$

where p_{nk} is the number of permutations of $\{1, \dots, n\}$ with k fixed elements.

- Show that the probability in (a) is asymptotic to $(1/2)e^{n-1}n!/n^{2n}$ as $n \rightarrow \infty$.

10281. *Proposed by Jonathan M. Borwein, University of Waterloo, Waterloo, Ontario, Canada.*

For $a > 0$ and $b > 0$, let

$$I(a, b) = \int_0^\infty \frac{t \, dt}{\sqrt[3]{(a^3 + t^3)(b^3 + t^3)^2}}.$$

(a) Show that

$$I(a, b) = I\left(\frac{a + 2b}{3}, \sqrt[3]{b \frac{a^2 + ab + b^2}{3}}\right).$$

(b) Show that the iteration which has $a_0 = a$ and $b_0 = b$ and

$$a_{n+1} = \frac{a_n + 2b_n}{3}$$

$$b_{n+1} = \sqrt[3]{b_n \frac{b_n^2 + a_n b_n + a_n^2}{3}}$$

converges to $I(1, 1)/I(a, b)$.

Notes: (10277) This produces an infinite family of counterexamples to a conjecture in a communication from Sophie Germain to Gauss on March 12, 1807. Details may be found in the article by N. MacKinnon, “Sophie Germain or Was Gauss a Feminist?,” *Math. Gazette*, 74 (1990), 346–351. **(10278)** This problem asks for the total number of admissible graphs, rather than for the number of admissible graphs which are incongruent under the group of isometries of the torus. **(10279)** This question arose in the doctoral dissertation of the author at New Mexico State University, Las Cruces, NM, and was formulated for use in this *Problem Section* with the help of W. H. Julian and C. Sweezy of that institution. **(10280)** A similar problem, where one restricts the definition of \star to the $n^{n(n+1)/2}$ commutative binary operations, may also be considered. Readers in need of more practice with these methods are encouraged to consider that as well. The axiom considered in this problem was found in a paper published in 1910 by Axel Thue on word problems in universal algebra (see Thue’s *Selected Mathematical Papers*, Oslo, 1977, 273–310). **(10281)** The problem is phrased to remind the reader of the arithmetic-geometric mean, as discussed in chapter 1 of J. M. Borwein and P. B. Borwein, *Pi and the AGM: Topics in Analytic Number Theory and Computational Complexity*, in the hope that a proof based on the form of the integral defining $I(a, b)$ will be found. The quantity $I(1, 1)$ appearing in part (b) is an elementary integral whose value is $2\pi\sqrt{3}/9$.

SOLUTIONS

Two Delicate Cyclic Inequalities

E3394 [1990, 529]. *Proposed by Mo Song-Qing, Institute of Applied Physics and Computational Mathematics, Beijing, China.*

Suppose $0 \leq a_1 \leq a_2 \leq \cdots \leq a_n$. For $2 \leq k \leq n-2$ consider the inequalities

$$\prod_{j=1}^n \frac{a_j + a_{j+1} + \cdots + a_{j+k-1}}{k} \leq \prod_{j=1}^n \frac{a_j + a_{j+1} + \cdots + a_{j+k}}{k+1}, \quad (1)$$

and

$$\sum_{j=1}^n (a_j a_{j+1} \cdots a_{j+k-1})^{1/k} \geq \sum_{j=1}^n (a_j a_{j+1} \cdots a_{j+k})^{1/(k+1)}, \quad (2)$$

where we adopt the convention that $a_{n+j} = a_j$.

(i) Prove that (1) and (2) hold for $k = 2$.

(ii)* Prove or disprove (1) and (2) for $3 \leq k \leq n-2$.

Solution by O. P. Lossers, Eindhoven University of Technology, Eindhoven, The Netherlands. We prove (i) and (ii), i.e., we show that both (1) and (2) hold for all k with $2 \leq k \leq n-2$. Our argument is divided into several parts. First we show that (1) and (2) are consequences of the theorem given below. Then we establish the lemma given below. Finally, we deduce the theorem from the lemma.

Recall that a doubly-stochastic matrix is a matrix with non-negative entries such that the sum of the entries in any row or column is equal to 1.

Theorem. Suppose $a_1 \leq a_2 \leq \cdots \leq a_n$ and $a_{n+j} = a_j$ for all j . Put

$$\alpha_{j,k} = k^{-1} \sum_{s=0}^{k-1} a_{j+s} \quad (2 \leq k \leq n-1).$$

Then for any fixed k with $2 \leq k \leq n-2$ there exists a doubly-stochastic matrix $M = (m_{ij})_{i,j=1}^n$ depending on $\{a_j\}_{j=1}^n$ such that

$$\alpha_{i,k+1} = \sum_{j=1}^n m_{ij} \alpha_{j,k} \quad (i = 1, 2, \dots, n).$$

Lemma. Suppose $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$, $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_n$, and $\sum_{i=1}^w \gamma_i \geq \sum_{j=1}^w \beta_j$ for $w = 1, 2, \dots, n$ with equality if $w = n$. Then there exists a doubly-stochastic matrix $M = (m_{ij})_{i,j=1}^n$ depending on $\beta_1, \dots, \beta_n, \gamma_1, \dots, \gamma_n$ such that

$$\gamma_i = \sum_{j=1}^n m_{ij} \beta_j \quad (i = 1, 2, \dots, n).$$

Part I. Deduction of (1) from the Theorem. In the notation of the theorem we must prove that

$$\prod_{i=1}^n \alpha_{i,k+1} \geq \prod_{j=1}^n \alpha_{j,k}.$$

We may assume that $\alpha_{j,k} > 0$ for all j . Using the theorem and the convexity of the exponential function we obtain

$$\alpha_{i,k+1} = \sum_{j=1}^n m_{ij} \exp(\ln \alpha_{j,k}) \geq \exp \sum_{j=1}^n m_{ij} (\ln \alpha_{j,k})$$

for $i = 1, 2, \dots, n$. Hence

$$\prod_{i=1}^n \alpha_{i,k+1} \geq \exp \sum_{i=1}^n \sum_{j=1}^n m_{ij} (\ln \alpha_{j,k}) = \exp \sum_{j=1}^n \ln \alpha_{j,k} = \prod_{j=1}^n \alpha_{j,k}.$$

Part II. Deduction of (2) from the Theorem. We must prove that if $0 \leq b_1 \leq b_2 \leq \dots \leq b_n$ and $b_{n+j} = b_j$ for all j , then

$$\sum_{i=1}^n (b_i b_{i+1} \cdots b_{i+k})^{1/(k+1)} \leq \sum_{i=1}^n (b_i b_{i+1} \cdots b_{i+k-1})^{1/k}.$$

If $b_1 = 0$, then

$$\begin{aligned} \sum_{i=1}^n (b_i b_{i+1} \cdots b_{i+k})^{1/(k+1)} &= \sum_{i=2}^{n-k} (b_i b_{i+1} \cdots b_{i+k})^{1/(k+1)} \\ &\leq \sum_{j=2}^{n-k} (b_{j+1} b_{j+2} \cdots b_{j+k})^{1/k} \\ &\leq \sum_{j=1}^n (b_j b_{j+1} \cdots b_{j+k-1})^{1/k}. \end{aligned}$$

If $b_1 > 0$, let us put $b_j = \exp a_j$ for $j = 1, 2, \dots, n$. Then by the theorem and the convexity of the exponential function we get

$$\exp \alpha_{i,k+1} = \exp \sum_{j=1}^n m_{ij} \alpha_{j,k} \leq \sum_{j=1}^n m_{ij} \exp \alpha_{j,k}.$$

Hence

$$\begin{aligned} \sum_{i=1}^n (b_i b_{i+1} \cdots b_{i+k})^{1/(k+1)} &= \sum_{i=1}^n \exp \alpha_{i,k+1} \\ &\leq \sum_{i=1}^n \sum_{j=1}^n m_{ij} \exp \alpha_{j,k} = \sum_{j=1}^n \exp \alpha_{j,k} \\ &= \sum_{j=1}^n \exp k^{-1}(a_j + a_{j+1} + \cdots + a_{j+k-1}) \\ &= \sum_{j=1}^n (b_j b_{j+1} \cdots b_{j+k-1})^{1/k}. \end{aligned}$$

Part III. Proof of The Lemma. For $n = 1$ the assertion is obvious. So assume that $m > 1$ and that the assertion of the lemma has been established for $n = m - 1$. Suppose $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_m$, $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_m$, and $\sum_{i=1}^w \gamma_i \geq \sum_{j=1}^w \beta_j$ for $w = 1, 2, \dots, m$ with equality if $w = m$. Since $\beta_1 \leq \gamma_1 \leq \beta_m$, it follows that for some $r \in \{1, 2, \dots, m-1\}$ we have $\gamma_1 \in [\beta_r, \beta_{r+1}]$, so that $\gamma_1 = \lambda \beta_r + (1 - \lambda) \beta_{r+1}$ for some $\lambda \in [0, 1]$. By the induction hypothesis there is an $m-1$ by $m-1$ doubly stochastic matrix V such that

$$\begin{pmatrix} \gamma_2 \\ \gamma_3 \\ \vdots \\ \gamma_m \end{pmatrix} = V \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_{r-1} \\ (1-\lambda)\beta_r + \lambda\beta_{r+1} \\ \beta_{r+2} \\ \vdots \\ \beta_m \end{pmatrix}.$$

We write V as a block matrix $[V_1 \ V_2 \ V_3]$, where V_1 , V_2 , and V_3 have $r-1$, 1 , and $m-r-1$ columns, respectively. Then it is easy to verify that the matrix M defined by

$$M = \begin{pmatrix} O_{r-1} & \lambda & 1-\lambda & O_{m-r-1} \\ V_1 & (1-\lambda)V_2 & \lambda V_2 & V_3 \end{pmatrix},$$

where O_{r-1} is a 1 by $r-1$ matrix of zeros and O_{m-r-1} is a 1 by $m-r-1$ matrix of zeros, has the required property of the lemma.

Part IV. Proof of the Theorem. Let $\beta_1 \leq \beta_2 \leq \dots \leq \beta_n$ and $\gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_n$ be non-decreasing orderings of $\alpha_{1,k}, \dots, \alpha_{n,k}$ and $\alpha_{1,k+1}, \dots, \alpha_{n,k+1}$ respectively. Since $a_1 \leq a_2 \leq \dots \leq a_n$, the sequence $\{\alpha_{m,k}\}_{m=1}^n$ is unimodal, namely, $\alpha_{1,k} \leq \alpha_{2,k} \leq \dots \leq \alpha_{n-k+1,k}$ and $\alpha_{n-k+1,k} \geq \alpha_{n-k+2,k} \geq \dots \geq \alpha_{n,k} \geq \alpha_{1,k}$. Thus for any integer w in $\{1, 2, \dots, n\}$ the maximum sum of any $n-w$ of the numbers $\alpha_{1,k}, \alpha_{2,k}, \dots, \alpha_{n,k}$ is attained for a sum of $n-w$ consecutive terms of $\{\alpha_{m,k}\}_{m=1}^n$, i.e., for a sum of the form $\sum_{j=t}^{t+n-w-1} \alpha_{j,k}$. Hence we have

$$\sum_{j=1}^w \beta_j = \sum_{j=1}^n \alpha_{j,k} - \max_t \sum_{j=t}^{t+n-w-1} \alpha_{j,k}$$

and similarly

$$\sum_{j=1}^w \gamma_j = \sum_{j=1}^n \alpha_{j,k+1} - \max_t \sum_{j=t}^{t+n-w-1} \alpha_{j,k+1}.$$

Hence

$$\begin{aligned} \sum_{j=1}^w \gamma_j - \sum_{j=1}^w \beta_j &= \max_t \sum_{j=t}^{t+n-w-1} \alpha_{j,k} - \max_t \sum_{j=t}^{t+n-w-1} \alpha_{j,k+1} \\ &= \max_t \sum_{j=t}^{t+n-w-1} k^{-1} \sum_{s=0}^{k-1} a_{j+s} - \max_t \sum_{j=t}^{t+n-w-1} (k+1)^{-1} \sum_{s=0}^k a_{j+s}. \end{aligned}$$

Reversing the order of summation we see that the first term here reduces to

$$\max_t k^{-1} \sum_{s=0}^{k-1} (n-w) \alpha_{t+s, n-w} = (n-w) \max_t k^{-1} \sum_{s=t}^{t+k-1} \alpha_{s, n-w}$$

and that the second term reduces to

$$\max_t (k+1)^{-1} \sum_{s=0}^k (n-w) \alpha_{t+s, n-w} = (n-w) \max_t (k+1)^{-1} \sum_{s=t}^{t+k} \alpha_{s, n-w}.$$

Since the sequence $\{\alpha_{s, n-w}\}_{s=1}^n$ is unimodal, a maximal arithmetic mean of k consecutive terms is at least as large as a maximal arithmetic mean of $k+1$ terms, i.e.,

$$\max_t k^{-1} \sum_{s=t}^{t+k-1} \alpha_{s, n-w} \geq \max_t (k+1)^{-1} \sum_{s=t}^{t+k} \alpha_{s, n-w}.$$

Hence $\sum_{j=1}^w \gamma_j \geq \sum_{j=1}^w \beta_j$ for $w = 1, 2, \dots, n$ with equality for $w = n$. Applying the lemma, we obtain the result of the theorem. Note that the doubly-stochastic matrix of the theorem is obtained from that of the lemma by suitable permutations of the rows and columns.

Editorial comment. If a_1, a_2, \dots, a_n are any non-negative real numbers, then both (1) and (2) hold for $k = 1$ and for $k = n - 1$; this can be readily seen by applying the inequality of the arithmetic and geometric means. However, for $2 \leq k \leq n - 2$ the monotonicity assumption $a_1 \leq a_2 \leq \dots \leq a_n$ is essential. Of course the assumption $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$ would serve as well.

Ingram Olkin and Larry Shepp proved (1) and (2) for $2 \leq k \leq n - 2$ by applying the theory of majorization. In fact they showed that (1) and (2) are only two of a large class of cyclic inequalities. The proposer supplied proofs of (1) and (2) for $k = 2$ and for $k = 3$.

The editors wish to thank Aimo Hinkkanen for his assistance.

No solutions other than those cited were received.

Balls and Urns at Random

E3412 [1990, 917]. *Proposed by Eric Wepsic, Boston, MA.*

Suppose that we place n balls (numbered from 1 to n) into n urns (numbered from 1 to n) in the following way: The i th ball is placed into an urn chosen randomly from the first i urns. Let $P(n, k)$ be the probability that an urn chosen at random from the n urns contains exactly k balls (i.e., $nP(n, k)$ is the expected number of urns containing exactly k balls). Find $\lim_{n \rightarrow \infty} P(n, k)$ for fixed k .

Composite solution by Peter Griffin, California State University, Sacramento, CA; Tim Hesterberg, Franklin and Marshall College, Lancaster, PA; and Richard Stong, University of California, Los Angeles, CA. The limit is 2^{-k-1} for $k = 0, 1, 2, \dots$. In fact we prove that for n a positive integer

$$P(n, 0) = \frac{1}{2} - \frac{1}{2n}, \quad (1)$$

$$P(n, 1) = \frac{1}{4} + \frac{1}{4n} + \frac{1}{2n^2}, \quad (2)$$

$$P(n, 2) = \frac{1}{8} + \frac{1}{8n} + \frac{1}{2n^2} \left(\sum_{j=1}^{n-1} \frac{1}{j} - \frac{1}{2} \right), \quad (3)$$

and we prove that for each $k \geq 1$ there exists a constant C_k such that

$$\left| P(n, k) - \frac{1}{2^{k+1}} - \frac{1}{2^{k+1}n} \right| \leq C_k n^{-2} \log^{k-1}(n+1) \quad (4)$$

for all positive integers n . Of course it is immediate from the definition of $P(n, k)$ that $P(j, k) = 0$ for $1 \leq j < k$ and $P(k, k) = 1/(k \cdot k!)$ for $k \geq 1$; both of these facts follow also by induction from the recursion formula (6) given below.

For $n \geq 2$, $k \geq 0$ we consider the result for n balls and n urns as being obtained from the result for $n - 1$ balls and $n - 1$ urns by adding another urn and another ball. An urn with k balls can then be obtained in one of four ways:

(i) the urn in question is one of the first $n - 1$ urns, it had k balls at the preceding stage, and the new ball is added to one of the $n - 1$ other urns;

(ii) the urn in question is one of the first $n - 1$ urns, it had $k - 1$ balls at the preceding stage, and it receives the new ball;

(iii) $k = 0$, the urn in question is the new urn, and the new ball is added to one of the first $n - 1$ urns;

(iv) $k = 1$, the urn in question is the new urn, and it receives the new ball.

These four possibilities lead to the four terms in the following recursion (for $n \geq 2$):

$$P(n, k) = \frac{n-1}{n}P(n-1, k)\frac{n-1}{n} + \frac{n-1}{n}P(n-1, k-1)\frac{1}{n} \\ + \delta_{k,0}\frac{1}{n}\frac{n-1}{n} + \delta_{k,1}\frac{1}{n}\frac{1}{n}.$$

This may be rewritten as

$$n^2P(n, k) - (n-1)^2P(n-1, k) \\ = (n-1)P(n-1, k-1) + (n-1)\delta_{k,0} + \delta_{k,1} \quad (5)$$

for $k \geq 0$ and $n \geq 2$. Here δ_{kl} is the Kronecker delta and it is understood that $P(n, -1) = 0$ for all positive integers n . In addition we obviously have the initial condition $P(1, k) = \delta_{k,1}$ for $k = 0, 1, 2, \dots$.

For $k = 0$ our recursion (5) becomes

$$n^2P(n, 0) - (n-1)^2P(n-1, 0) = n-1 \quad (n \geq 2),$$

with the initial condition $P(1, 0) = 0$. Summation gives

$$n^2P(n, 0) - P(1, 0) = \sum_{j=2}^n (j-1) = n(n-1)/2.$$

Thus (1) follows.

For $k = 1$ the recursion (5) becomes

$$n^2P(n, 1) - (n-1)^2P(n-1, 1) = 1 + (n-1)P(n-1, 0) \quad (n \geq 2),$$

with the initial condition $P(1, 1) = 1$. Summation gives

$$n^2P(n, 1) - P(1, 1) = \sum_{j=2}^n (1 + (j-1)P(j-1, 0)) = \sum_{j=2}^n \frac{j}{2}.$$

Thus (2) follows.

For $k \geq 2$ the recursion (5) becomes

$$n^2P(n, k) - (n-1)^2P(n-1, k) = (n-1)P(n-1, k-1) \quad (n \geq 2),$$

with the initial condition $P(1, k) = 0$. Summation gives

$$n^2P(n, k) = \sum_{j=2}^n (j-1)P(j-1, k-1) \quad (6)$$

for $k \geq 2$, $n \geq 2$. In particular, a brief calculation for $k = 2$ gives (3), so that

$$P(n, 2) = \frac{1}{8} + \frac{1}{8n} + \frac{\log n + \gamma - \frac{1}{2} + O(n^{-1})}{2n^2},$$

where γ is Euler's constant. Consequently (4) holds for $k = 2$. To complete the proof of (4), we suppose $k \geq 3$ and we make the inductive assumption that

$$\left| P(n, k-1) - \frac{1}{2^k} - \frac{1}{2^k n} \right| \leq C_{k-1} n^{-2} \log^{k-2}(n+1)$$

for all positive integers n . Now (6) gives

$$\begin{aligned} n^2 \left(P(n, k) - 2^{-k-1} - \frac{2^{-k-1}}{n} \right) \\ = \sum_{j=2}^n (j-1) \left(P(j-1, k-1) - 2^{-k} - \frac{2^{-k}}{j-1} \right) - 2^{-k}, \end{aligned}$$

so that by our inductive assumption

$$\begin{aligned} n^2 \left| P(n, k) - 2^{-k-1} - \frac{2^{-k-1}}{n} \right| &\leq \sum_{j=2}^n C_{k-1} (j-1)^{-1} \log^{k-2} j + 2^{-k} \\ &\leq C_k \log^{k-1} (n+1) \end{aligned}$$

for a suitable choice of C_k . Thus (4) is proved.

The formula (6) can be used to calculate the numerical values of $P(n, k)$. Since $P(j, k) = 0$ for $1 \leq j < k$ and $P(k, k) = 1/(k \cdot k!)$ for $k \geq 1$, (6) may be rewritten as

$$\begin{aligned} n^2 P(n, k) &= \sum_{j=k}^n (j-1) P(j-1, k-1) = 1/(k-1)! \\ &+ \sum_{j=k+1}^n (j-1) P(j-1, k-1) \quad (n \geq k \geq 2). \end{aligned}$$

We remark that for any fixed $k \geq 2$ a slight amplification of the above argument will replace (4) by the stronger result

$$P(n, k) = \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}n} + \frac{(\log n + \gamma - 1/2)^{k-1}}{2(k-1)!n^2} + O\left(\frac{\log^{k-3}(n+1)}{n^2}\right) \quad (7)$$

or even by the much stronger result

$$P(n, k) = \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}n} + \frac{p_k(\log n)}{2(k-1)!n^2} + O\left(\frac{\log^{k-2}(n+1)}{n^3}\right), \quad (8)$$

where p_k is a certain monic polynomial of degree $k-1$ whose coefficient of $(\log n)^{k-2}$ is $(k-1)(\gamma - 1/2)$.

Editorial comment. Lajos Takács remarked that $P(n, k)$ may be expressed in terms of the unsigned Stirling numbers of the first kind $S(n, k)$ defined for $n \geq k \geq 1$ by the expansion

$$\sum_{k=1}^n S(n, k) x^k = x(x+1)(x+2) \cdots (x+n-1).$$

Specifically, Takács observed that for $k \geq 1$ we have

$$\begin{aligned} n!nP(n, k) &= \frac{1}{2}S(n, k) + 2^{-k-1} \sum_{j=k}^n 2^j S(n, j) \\ &= 2^{-k-1}(n+1)! + \frac{1}{2}S(n, k) - 2^{-k-1} \sum_{j=1}^{k-1} 2^j S(n, j). \end{aligned} \quad (9)$$

Now let σ_k be the k th elementary symmetric function of the $n-1$ numbers

$1, 1/2, 1/3, \dots, 1/(n-1)$ and let s_k be the sum of their k th powers

$$s_k = \sum_{j=1}^{n-1} \frac{1}{j^k}.$$

Since

$$\sum_{k=1}^n S(n, k) x^k = (n-1)! x \prod_{j=1}^{n-1} \left(1 + \frac{x}{j}\right),$$

we see that $S(n, 1) = (n-1)!$ and

$$S(n, k) = (n-1)! \sigma_{k-1} \quad (n \geq k > 1). \quad (10)$$

Now by using Newton's formula

$$k\sigma_k = s_1\sigma_{k-1} - s_2\sigma_{k-2} + \dots + (-1)^{k-2}s_{k-1}\sigma_1 + (-1)^{k-2}s_k$$

and induction on k it is easy to see that

$$k!\sigma_k = s_1^k + \sum_{j=0}^{k-2} s_1^j P_{kj}(s_2, \dots, s_k), \quad (11)$$

where P_{kj} is a polynomial in $k-1$ variables with integer coefficients. (For example $\sigma_1 = s_1$, $2\sigma_2 = s_1^2 - s_2$, $6\sigma_3 = s_1^3 - 3s_1s_2 + 2s_3$, $24\sigma_4 = s_1^4 - 6s_1^2s_2 + 8s_1s_3 + 3s_2^2 - 6s_4$). Since $s_1 = \log n + \gamma + O(1/n)$ and $s_r = \zeta(r) + O(n^{1-r})$ for $r \geq 2$, we readily deduce from (10) and (11) that for $k \geq 2$ we have

$$(k-1)!S(n, k) = (n-1)!((\log n + \gamma)^{k-1} + g_k(\log n + \gamma) + O(n^{-1} \log^{k-2}(n+1))), \quad (12)$$

where g_k is a certain polynomial in one variable of degree $k-3$ (g_2 being identically zero).

The papers, C. Jordan, "On Stirling's numbers," *Tôhoku Math. J.* 37 (1933), 254-278, R. Jungen, "Sur les séries de Taylor n'ayant que des singularités algébriques-logarithmiques sur leur cercle de convergence", *Comm. Math. Helv.* 3 (1931), 266-306, and L. Moser and M. Wyman, "Asymptotic development of the Stirling numbers of the first kind," *J. London Math. Soc.* 33 (1958), 133-146, should be consulted for further details.

Solved also by A. Adler, R. A. Agnew, D. Callan, E. Hertz, R. High, R. D. Hurwitz, O. P. Lossers (The Netherlands), J. McHugh, M. F. Neuts & Q. M. He, F. Richman, A. J. Rosenthal, Anchorage Math Solutions Group, Con Amore Problem Group (Denmark), National Security Agency Problems Group, and the proposer. One incorrect solution was received.

Coefficients of a Generating Function

E3415 [1991, 54]. *Proposed by Philippe Flajolet and Donald E. Knuth, Stanford University, Stanford, CA.*

Find the coefficient of $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ in

$$(1-x_1)^{-a_1} (1-x_1-x_2)^{-a_2} \cdots (1-x_1-x_2-\cdots-x_n)^{-a_n}.$$

Solution by Dean Alvis, Indiana University, South Bend, IN. For $1 \leq j \leq n$, let $s_j = a_j + k_j$. Then the desired coefficient may be expressed as

$$\binom{s_n-1}{k_n} \binom{s_n+s_{n-1}-1}{k_{n-1}} \cdots \binom{s_n+\cdots+s_1-1}{k_1}.$$

To see this, let

$$f_{a_1, a_2, \dots, a_n}(x_1, \dots, x_n) = (1 - x_1)^{-a_1} (1 - x_1 - x_2)^{-a_2} \cdots (1 - x_1 - x_2 - \cdots - x_n)^{-a_n}.$$

The basis for the induction will be the case $n = 1$. The coefficient of x^k in $f_a(x) = (1 - x)^{-a}$ is $\binom{a+k-1}{k}$ and may be obtained by using the generalized binomial theorem or as a special case of the formula given below for the inductive step. We turn to the inductive step. Here we have the differentiation formula

$$\begin{aligned} & \left. \frac{\partial^{k_n}}{\partial x_n^{k_n}} f_{a_1, a_2, \dots, a_n}(x_1, \dots, x_n) \right|_{x_n=0} \\ &= a_n(a_n + 1) \cdots (a_n + k_n - 1) (1 - x_1)^{-a_1} (1 - x_1 - x_2)^{-a_2} \\ & \quad \cdots (1 - x_1 - \cdots - x_{n-1})^{-(a_{n-1} + a_n + k_n)} \\ &= k_n! \binom{s_n - 1}{k_n} f_{a_1, a_2, \dots, a_{n-1} + s_n}(x_1, \dots, x_{n-1}). \end{aligned}$$

Thus, the coefficient of $x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ in $f_{a_1, a_2, \dots, a_n}(x_1, \dots, x_n)$ is equal to $\binom{s_n - 1}{k_n}$ times the coefficient of $x_1^{k_1} x_2^{k_2} \cdots x_{n-1}^{k_{n-1}}$ in $f_{a_1, a_2, \dots, a_{n-1} + s_n}(x_1, \dots, x_{n-1})$.

Therefore, the answer follows by induction on n .

Editorial comment. Most solutions received were similar to that given above except that the role of partial derivatives was replaced by the use of $(1 - x)^{-a} = \sum_{j=0}^{\infty} \binom{a+j-1}{j} x^j$. O. P. Lossers showed that the induction argument could also be done easily by eliminating the variable x_1 rather than x_n . Daniel Brown obtained the answer in the case that the a_j 's are non-negative integers from a simple combinatorial interpretation of the problem. Volker Strehl and the proposers offered independent generalizations of the problem.

Solved also by S.-J. Bang (Korea), J. C. Binz (Switzerland), D. Brown (Canada), D. Callan, H. Lipman, O. P. Lossers (The Netherlands), A. Nijenhuis, R. Stong, V. Strehl (Germany), R. Will, J. Zenq (France), and the proposers. The Anchorage Math Solutions Group gave the correct answer but no details of the proof. Another reader used a method of proof which led to an incorrect answer.

Characterization of an Iterable Function

E3428 [1991, 263]. *Proposed by Artin B. Boghossian, Aramex Investments Inc., Willowdale, Ontario, Canada.*

Let S be a non-empty interval on the real line. Let $f: S \rightarrow S$ be a continuous function having the property that for each $x \in S$ there exists a positive integer $n = n(x)$ with $f^n(x) = x$, where f^n denotes the n th iterate of f . For given S characterize all such functions.

Solution by Dan Velleman, Amherst College, Amherst, MA. Let us call functions of the required kind *iterable*. Clearly, the identity function is iterable. We show that any other iterable function f must be a continuous decreasing bijection from S to itself such that f^2 is the identity. The graph of such a function is symmetric about the line $y = x$. If S contains both or neither of its endpoints, there are many such functions; otherwise, the only iterable function is the identity.

Suppose f is iterable. Clearly, f is onto. To see that f is one-to-one, suppose that $f(a) = f(b)$ and choose positive integers m and n such that $f^m(a) = a$ and $f^n(b) = b$. Then $a = f^{mn}(a) = f^{nm}(b) = b$. Hence f is a bijection and by the Intermediate Value Theorem it follows that f is monotone.

If f is increasing, then f must be the identity. For, if $f(x) > x$, it follows that $f^{n+1}(x) > f^n(x)$ for all $n \in \mathbb{N}$. Thus, $f^n(x) > x$ for all $n > 0$, so f cannot be iterable. A similar argument applies if $f(x) < x$.

If f is decreasing and $f^n(x) = x$, then $(f^2)^n(x) = (f^n)^2(x) = x$. Hence f^2 is an increasing iterable function, which we have just seen must be the identity.

Editorial comment. Most respondents correctly showed that these functions are precisely those whose square is the identity, but overlooked the role of the endpoints in determining the number of these functions. N. P. Bhatia and W. O. Egerland pointed out that the crux of the matter is that the interval spanned by an n -periodic orbit of a continuous function contains a non-periodic point if $n \geq 3$. R. High showed that the “pointwise periodicity” may be relaxed to “pointwise almost periodicity” in the following sense: for each $x \in S$ and each $\varepsilon > 0$, there exists an n for which $f^n(x)$ is within ε units of x . High also observed that, if S is the unit circle, then there is a richer set of solutions to the “pointwise periodic” problem, including rotations through a rational angle. In that setting, solutions to the analogous “pointwise almost periodic” problem further include rotations through an irrational angle. David Callan mentioned the related problems No. 993 in *Mathematics Magazine* [1976, 212; 1978, 130] and No. 6133 in this MONTHLY [1977, 140; 1978, 771].

Solved also by 25 readers (including those cited) and the proposer. One incorrect and four incomplete solutions were also received.

A Sequence Converging to Zero

E3441 [1991, 438]. *Proposed by Xu Chenglong, Shanghai University of Science and Technology, China.*

Suppose $0 \leq p < 1$ and $q \geq 0$. Put

$$U_n = \sum_{k=0}^n \binom{n}{k} p^{n-k} q^k \frac{1}{k!}.$$

Determine $\lim_{n \rightarrow \infty} U_n$.

Solution I by H. Turner Laquer, Idaho State University, Pocatello, ID. The limit is 0. This follows from the fact that $\sum_{n=0}^{\infty} U_n$ is bounded. Indeed,

$$\begin{aligned} \sum_{n=0}^{\infty} U_n &= \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} p^{n-k} q^k \frac{1}{k!} = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \binom{n}{k} p^{n-k} q^k \frac{1}{k!} \\ &= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{k!} q^k \binom{k+j}{k} p^j = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{q^k}{(1-p)^{k+1}} = \frac{1}{1-p} e^{q/(1-p)}. \end{aligned}$$

Solution II by Eugene A. Herman, Grinnell College, Grinnell, IA. The limit is 0 for all complex numbers p and q with $|p| < 1$. If $p = 0$, then $U_n = q^n/n! \rightarrow 0$ for all q . If $p \neq 0$, choose $\alpha > 0$ such that $1 + \alpha < |p|^{-1}$, which is possible since $|p| < 1$. Since $q^n/n!$ has a limit, its magnitude is bounded, and similarly

$|(q/p\alpha)^k/k!|$ is bounded by some value M . Then

$$|U_n| = |p^n| \left| \sum_{k=0}^n \binom{n}{k} \alpha^k \left(\frac{q}{p\alpha} \right)^k \frac{1}{k!} \right| \leq M |p^n| \sum_{k=0}^n \binom{n}{k} \alpha^k = M |p(1+\alpha)|^n,$$

which has limit 0 since $|p(1+\alpha)| < 1$.

Solution III by Victor Hernández, Universidad Nacional de Educación a Distancia, Madrid, Spain. The limit is 0. Letting $r = q/(1-p)$, we have

$$e^{-r}U_n = \sum_{k=0}^n p^{n-k}(1-p)^k(e^{-r}r^k/k!).$$

Hence $e^{-r}U_n$ equals $P(X=Y)$, where X is a binomial random variable with parameters n and $(1-p)$, Y is a Poisson random variable with fixed parameter r , and X, Y are independent. As $n \rightarrow \infty$, $P(X < n(1-p)/2) \rightarrow 0$ and $P(Y \geq n(1-p)/2) \rightarrow 0$, so $P(X=Y) \rightarrow 0$.

Editorial comment. The method of Solution I was the most popular. It may also be extended to prove the result stated in Solution II. Other worthy solutions used yet other methods. B. E. Rhoades, David Borwein, and Ignacy Icchak Kotlarski noted that $\langle U_n \rangle$ is a transform of a null sequence, the first two seeing it as an Euler transform and the third as an inverse Laplace transform. The solution by Hongzhu Qiao and Riuming Zhang and the solution by David Callan related an expression involving U_n to Laguerre polynomials. Kenneth F. Andersen observed that U_n is the value at $x = 0$ of the n th derivative of $e^{px}I_0(2\sqrt{qx})$, where I_0 is the modified Bessel function. Not to be outdone, Terence R. Shore and Douglas B. Tyler provided six proofs, including a bound on the rate of convergence of a more general sequence. If $p < \beta < 1$ and $\sum_{k=0}^{\infty} a_k z^k$ is any entire function, they proved there exists a constant $K > 0$ such that $\sum_{k=0}^n \binom{n}{k} p^{n-k} a_k \leq K\beta^n$ for all $n \geq 0$.

Solved by 36 readers, including those cited, and the proposer.

The sphere from off-center

E3460 [1991, 755]. *Proposed by E. Ehrhart, University of Strasbourg, France.*

(a) Suppose we have n mutually perpendicular chords through a point P interior to a sphere S in n -dimensional Euclidean space. Prove that the sum of the squares of the lengths of these chords depends only on the radius r of the sphere and the distance d from P to the center of the sphere.

(b) More generally, suppose $1 \leq k \leq n$. Each set of k of the n mutually perpendicular chords through P given in (a) determines a k -dimensional affine subspace. Prove that the sum of the $(2/k)$ -th powers of the k -dimensional measure of the cross-sections of S made by these $\binom{n}{k}$ affine subspaces depends only on r and d .

Solution by Albert Nijenhuis, Seattle, WA. The value is: (a) $4(nr^2 - (n-1)d^2)$; (b) $\alpha_k^{2/k} \left[\binom{n}{k} r^2 - \binom{n-1}{k} d^2 \right]$, where α_k is the volume of the unit k -ball.

Choose coordinates along the chords, so P is the origin; let \vec{c} be the vector from P to the center and \vec{x} the running coordinates. Then the equation of S is $(\vec{x} - \vec{c}) \cdot (\vec{x} - \vec{c}) = r^2$. Further, $\vec{c} \cdot \vec{c} = d^2$.

(a) The coordinates of the endpoints of the i -th chord are the solutions of the quadratic equation $x^2 - 2c_i x + d^2 - r^2 = 0$, so the square of the length of the i -th chord is

$$\left[\frac{2c_i + \sqrt{4c_i^2 - 4(d^2 - r^2)}}{2} - \frac{2c_i - \sqrt{4c_i^2 - 4(d^2 - r^2)}}{2} \right]^2$$

$$= 4c_i^2 - 4(d^2 - r^2).$$

Summation on i gives $\sum_{i=1}^n 4(c_i^2 - d^2 + r^2) = 4(nr^2 - (n-1)d^2)$.

(b) Let $\vec{\gamma}$ be the vector of the first k components of \vec{c} , and \vec{y} the running coordinates in the subspace spanned by the first k axes. Then S meets this subspace in a $k-1$ -sphere, whose equation is $\vec{y} \cdot \vec{y} - 2\vec{\gamma} \cdot \vec{y} + d^2 - r^2 = 0$. Writing the equation as

$$(\vec{y} - \vec{\gamma}) \cdot (\vec{y} - \vec{\gamma}) - \vec{\gamma} \cdot \vec{\gamma} + d^2 - r^2 = 0,$$

we see that the square of its radius is $r^2 - d^2 + \vec{\gamma} \cdot \vec{\gamma}$. The volume of the k -ball bounded by this sphere is $\alpha_k (r^2 - d^2 + \vec{\gamma} \cdot \vec{\gamma})^{k/2}$. Similar expressions are obtained for the remaining selections of k chords. Summing the $(2/k)$ -th powers of all $\binom{n}{k}$ volumes we obtain

$$\alpha_k^{2/k} \sum (r^2 - d^2 + \vec{\gamma} \cdot \vec{\gamma}) = \alpha_k^{2/k} \left[\binom{n}{k} (r^2 - d^2) + \sum \vec{\gamma} \cdot \vec{\gamma} \right].$$

Now $\vec{\gamma}$ denotes the vector to the center of each one of the $\binom{n}{k}$ k -balls in turn. In the sum $\sum \vec{\gamma} \cdot \vec{\gamma}$ the term c_i^2 occurs as often as the i -th axis is one of a set of k axes, i.e., $\binom{n-1}{k-1}$ times. Therefore, $\sum \vec{\gamma} \cdot \vec{\gamma} = \binom{n-1}{k-1} d^2$, and the sought-for sum equals

$$\alpha_k^{2/k} \left[\binom{n}{k} (r^2 - d^2) + \binom{n-1}{k-1} d^2 \right] = \alpha_k^{2/k} \left[\binom{n}{k} r^2 - \binom{n-1}{k} d^2 \right].$$

Solved also by D. W. Bailey, D. Batman, R. J. Chapman (U.K.), P. Čížek (student, Czechoslovakia), E. Dahlman, I. Dimitric, M. Golomb, H. Kappus (Switzerland), I. Kastanas, K. S. Kedlaya (student), N. Komanda, O. P. Lossers (The Netherlands), D. Magagnosc, K. Schilling, J. B. Wilker (Canada), Anchorage Math Solutions Group, and the proposer.

Collaborating editors: David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pfiefer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, Douglas B. Tyler, Daniel Ullman, Edward T. H. Wang, and William E. Watkins.

More Evidence for Bourbaki

For a number of years, the MONTHLY has been engaged in a debate about the existence of the mathematician Bourbaki. Recent evidence was sent to us all the way from Paris, France, by Jean-Pierre Grivaux. It is nothing less than a Plate of Bourbaki himself. Unfortunately, it is not a *photographic* plate, but rather a table plate, with the monogram CB. Grivaux writes that it was given as a reward to his great-grandfather by the unfortunate General *Charles Bourbaki* during the Franco-Prussian war in 1871. Alas, the precise relationship of Charles to Nicolas (the mathematician) is not known. The MONTHLY remains unconvinced.



The American Mathematical Monthly



Volume 100, Number 2 / FEBRUARY 1993



AN OFFICIAL PUBLICATION OF THE MATHEMATICAL ASSOCIATION OF AMERICA

NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International,
Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 2 / FEBRUARY 1993
(ISSN 0002-9890)



Contents

ARTICLES

Yueh-Gin Gung and Dr. Charles Y. Hu Award for Distinguished Service to
Henry O. Pollak / RAMANATHAN GNANADESIKAN
and HENRY J. LANDAU 115

John Marvin Colaw: and The American Mathematical Monthly /
JOHN D. MAXWELL 117

Newton and the Transmutation of Force / TRISTAN NEEDHAM 119

A Note on Diophantine Representations / CHRISTOPH BAXA 138

Recurrence of Simple Random Walk in the Plane /
TERENCE R. SHORE and DOUGLAS B. TYLER 144

Pick's Theorem / BRANKO GRÜNBAUM AND G. C. SHEPHARD 150

Mathematics for Liberal Arts Students / ALBERT W. BRIGGS, JR. 162

FEATURES

COMMENTS 114

NOTES 167

PICTURE PUZZLE 175

THE AUTHORS 176

LETTERS 178

UNSOLVED PROBLEMS

A mod- n Ackermann Function, or What's So Special About 1969? /
JON FROEMKE AND JERROLD W. GROSSMAN 180

PROBLEMS AND SOLUTIONS 184

REVIEWS

How to Read and Do Proofs, by Daniel Solow / M. F. JANOWITZ 197

TELEGRAPHIC REVIEWS 200

Newton and the Transmutation of Force

Tristan Needham

1. INTRODUCTION. The year 1687 witnessed an event that was to dramatically alter the course of science, indeed, of western civilization—the publication of Sir Isaac Newton’s *Philosophiae Naturalis Principia Mathematica* [1], better known simply as the *Principia*. Ten years ago, having been intrigued by the mathematical glimpses contained in Westfall’s magnificent biography [2], I took up in earnest the study of Newton’s great work, and the experience was as dazzling as anything one might hope to encounter on the road to Damascus.

The primary purpose of the paper is to present a new geometric way of understanding a beautiful but little-known fact concerning the transmutation of central force fields by means of complex mappings. A second purpose, however, is to swell the Newtonian congregation! To this end, in this introduction we shall supply five very elementary examples (suitable for the classroom) of the power of the *Principia*’s methods. This should also help to eliminate the potential ‘culture shock’ presented by the chosen methodology of the rest of the paper. First, though, a few general remarks on the *Principia* are perhaps in order. Since no mention will be made in this section of the principal problem to be analyzed, the impatient reader is welcome to jump to the next section where the work itself is begun.

As illustrated in [3], one motivation for the study of the *Principia* is that *unknown results* of importance to modern mathematics may thereby be revealed. For example, in his splendid little book on seventeenth century science [4], V. I. Arnol’d devotes an entire chapter to the *Principia*’s astonishing but neglected Lemma XXVIII ([1], p. 110). As he explains, this is in fact a brilliant topological proof of a result on the transcendence of Abelian integrals!

No less exciting than Newton’s results, is the *method* he used to obtain them. Arnol’d puts it well: referring to Prop. LXX ([1], p. 193), “This sample of Newton’s argument shows how it is possible to solve problems from potential theory without analysis, without knowing the theory of harmonic functions, or the fundamental solution of the Laplace equation, or the simple and double layer potentials. Similar arguments, preceding the rise of analysis, often occur in papers of those times and turn out to be very powerful” ([4], p. 27). The expression “preceding the rise of analysis” is rather curious in this context, for Newton had himself invented the calculus some twenty years before constructing the geometric argument to which Arnol’d refers. This observation leads us to the following famous ‘paradox.’

Why did the inventor of the calculus not use it in the *Principia*? The answer—according to a persistent and pernicious myth—is that he *did* use calculus to make his great discoveries, but then disguised the fact (for Machiavelian reasons) by translating his arguments into geometry. In the hope that the wide readership of the *Monthly* may help to put this three-century old story to rest, let

No such papers demonstrating propositions of the *Principia* in a form different from that published have ever been found, except for a few in which he later set a couple of propositions over into analytical terms. The problem vanishes, however, when we view the *Principia* against the background of Newton's mathematical development in the years immediately preceding. Around 1680, the study of ancient geometry led Newton to a revulsion from the inelegant demonstrations of modern analysis.

(1) Let $T = \tan \theta$ and suppose we wish to explain why $dT/d\theta = 1 + T^2$. FIG. 1 illustrates the increase $\Delta T (= cr)$ in T that results from an increase of $\Delta\theta$ in θ . With a as center, draw the circle through c , and let the tangent to this circle at c cut ar in s , thereby producing the equal angles cab and rsc . The desired result follows almost instantly upon observing the behavior of the triangle crs in the limit as $\Delta\theta \rightsquigarrow 0$. Already in FIG. 1 it is hard to distinguish between cs and the arc $L \cdot \Delta\theta$, and indeed at the very end of the shrinking process their "last ratio" (as Newton would say) will be unity. Thus,

But in the shrinking process it is also clear that the triangle crs becomes similar to acb , and hence an alternative expression to the above is

120

Combining these two views of the limit process, we conclude that

$$\frac{dT}{d\theta} = L^2 = 1 + T^2.$$

In the firm belief that the most persuasive advocate of the Newtonian approach will be you yourself, we present the remaining examples as problems.

(2) Reconsider FIG. 1. Instead of looking at triangles with constant unit base, consider those of constant unit hypotenuse. By drawing the diagram for the new case, deduce that $(\sin \theta)' = \cos \theta$ and $(\cos \theta)' = -\sin \theta$. Note the economy with which one diagram yields both results.

(3) In the xy -plane, consider the family of triangles bounded by the coordinate axes and a variable line L through the fixed point $Q(a, b)$. See FIG. 2(A). By drawing the change in area resulting from a tiny rotation of L about Q , find the minimum area of such a triangle. Note the generality of the Newtonian reasoning. For example, consider the shaded area in FIG. 2(B) that is cut off from the ellipse by L as it rotates about the arbitrary interior point Q . Deduce that the area will be minimum [maximum, upon rotation of L by π] when L is parallel to the tangent at P .

(4) Consider FIG. 2(C). A basic property of ellipses that we shall need later in the paper is that the sum of the focal distances (F_1P and F_2P) is constant, and hence equal to the major axis. By drawing a diagram of the changes in the focal distances that result when P is minutely displaced along the ellipse, deduce the ‘reflection property’: light emitted from F_1 is reflected to F_2 . In like manner, deduce the corresponding reflection properties of the other two conic sections.

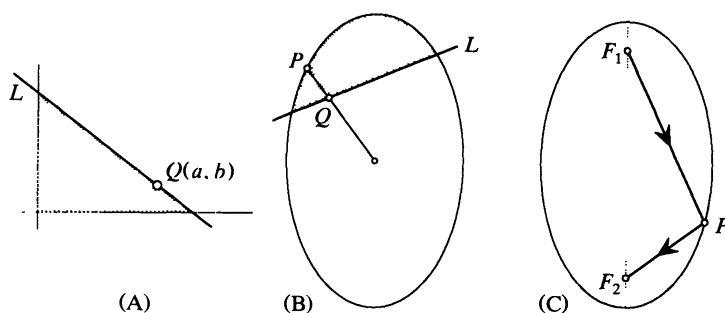


Figure 2.

(5) Arnol'd [4], p. 28: “Here is an example of a problem that people like Barrow, Newton and Huygens would have solved in a few minutes and which present-day mathematicians are not, in my opinion, capable of solving quickly [The only exception I know—G. Faltings—proves the rule.]: to calculate

$$\lim_{x \rightarrow 0} \frac{\sin(\tan x) - \tan(\sin x)}{\arcsin(\arctan x) - \arctan(\arcsin x)}.”$$

We shall add a hint that Arnol'd does not supply: think what the graphs of the four functions must look like (*roughly*) near the origin. The answer is given on p. 108 of [4].

Certainly these problems were not chosen at random, but I trust that a point has nevertheless been made. Having been sensitized to the possibility of such solutions, you will no doubt find many other examples of your own whenever you next have occasion to teach calculus.

Although I derived much pleasure and knowledge from the application of the *Principia*'s ideas to problems of ordinary calculus, the real reward came with the gradual realization that these ideas could be applied, very naturally, to *complex* analysis, yielding an approach that I hold to be considerably more elementary and intuitive than the conventional one. These ideas have now taken on the definite form of a forthcoming book from Oxford University Press, entitled *Visual Complex Analysis*. The following is essentially an expanded extract from that work.

2. THE TALE OF THE POUND NOTE. In commemoration of the three-hundredth anniversary of the publication of the *Principia* in 1687, the Bank of England brought out the new pound note shown in FIG. 3. The diagram that it bears is indeed fitting, for it is a faithful reproduction of that found in Prop. XI ([1], p. 56), and it represents Newton's solution to the problem that catalyzed his entire endeavor in the *Principia*: to mathematically link the observed motion of the planets—which Kepler had found to be elliptical—to a solar attraction diminishing as the square of the distance.

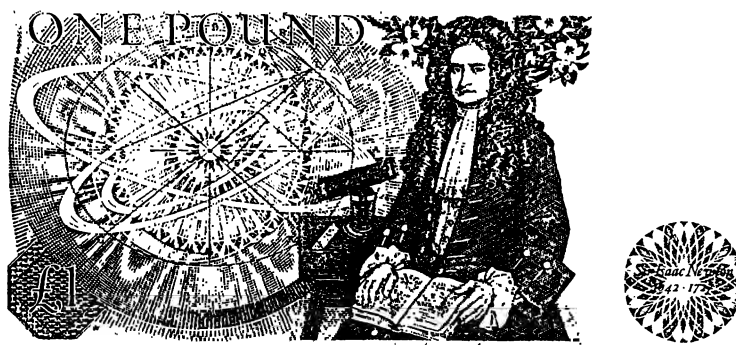


Figure 3.

Unhappy (presumably) with the baldness of Newton's figure, the decision was taken to embellish the pound note with what is clearly intended to represent the sun and some attendant bodies in orbit. But in doing so an extraordinary blunder was committed, for instead of being at the focus of the ellipse, this sun sits squarely at its *center*! Here our tale begins, for a closer examination of the *Principia* will reveal that there exists a point of view from which this blunder ceases to be a blunder at all.

Early in his investigations, Newton realized—even leaving aside their physical importance—that two special power laws enjoyed a degree of mathematical elegance not shared by the others: force decreasing as the square of the distance, and force increasing directly as the distance. He further observed that these “two principal cases” (as he called them) exhibited striking similarities. For example, only for these two laws are the orbits conic sections. This fact inspired Newton to go further.

Immediately following the direct proof employing FIG. 3, he gave an ingenious alternative argument in which he showed that the inverse square law for the force directed to the sun at the focus is a *consequence* of the linear force law that would be required if the elliptical orbit were instead caused by a fictitious attracting body at the center. In the light of this fact, the artist's embellishment no longer seems out of place!

Newton couched his argument in terms of a more general result which we will state now and prove later. Consider FIG. 4. A force field F_{OLD} emanating from A holds P in a given orbit. Newton asks the question, if the center of attraction is moved from A to any other place B , then into what new force field F_{NEW} must F_{OLD} be transmuted in order that the orbit of P remain the same? Draw AQ parallel to BP and meeting the tangent at P in Q . Newton's answer (Cor. III, Prop. VII) is

$$F_{\text{NEW}} \propto \left(\frac{AQ^3}{BP^2 \cdot AP} \right) F_{\text{OLD}}. \tag{1}$$

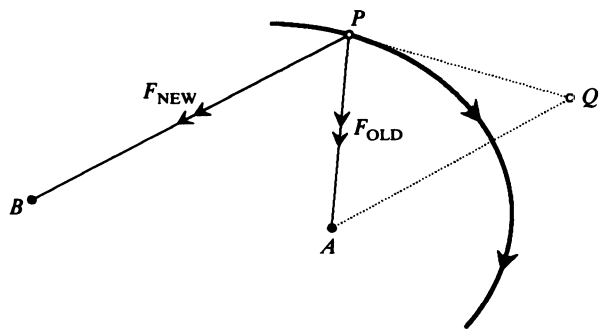


Figure 4.

Let us now follow Newton's beautiful application of this result to the transmutation of a linear force field into an inverse square field. Consider FIG. 5, in which the body at P is shown orbiting the sun at the focus B , and in which SC and RA

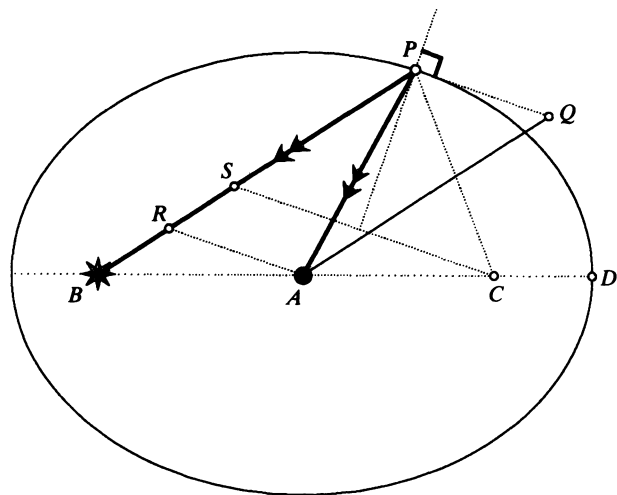


Figure 5.

have been drawn parallel to the tangent PQ . We are now to think of the gravitational field emanating from the focus as arising out of the field that a fictitious body at the center would need to possess in order to hold P in its given elliptical orbit:

$$F_{\text{FOCUS}} \propto \left(\frac{AQ^3}{BP^2 \cdot AP} \right) F_{\text{CENTER}}. \quad (2)$$

Notice two things: $BA = CA$ implies $BR = SR$; the reflection property established in the introduction says that BP and CP make equal angles with the normal at P , and thus $PS = CP$. It follows that

$$AQ = PR = \frac{BP + PS}{2} = \frac{BP + CP}{2} = AD,$$

and since AD is constant, (2) therefore becomes

$$F_{\text{FOCUS}} \propto \left(\frac{1}{BP^2 \cdot AP} \right) F_{\text{CENTER}}.$$

But Newton has already established in Prop. X that $F_{\text{CENTER}} \propto AP$, and he thereby concludes that the gravitational field of the sun is

$$F_{\text{FOCUS}} \propto \frac{1}{BP^2}.$$

3. ENTER THE COMPLEX PLANE. Consider a particle of unit mass located at the point z in the complex plane, and subject to a force $|z|$ directed towards the origin. The differential equation governing its motion will therefore be $\ddot{z} = -z$, the general solution of which is

$$z = pe^{it} + qe^{-it}, \quad (3)$$

where p and q may, without any real loss of generality, be taken as real and satisfying $p > q$. As illustrated on the left of FIG. 6, this is elliptical motion with the attracting point at the center, and with foci at $\pm 2\sqrt{pq}$. These facts are

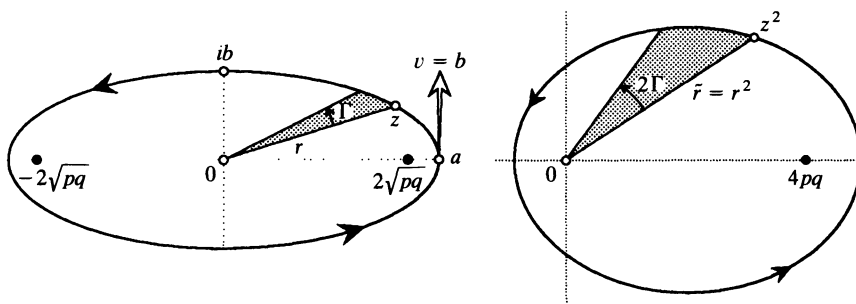


Figure 6.

perhaps more readily verified upon re-expressing (3) as $z = a \cos t + ib \sin t$, where $a = p + q$ and $b = p - q$. Of course each of these numbers has a double significance: a is both the semimajor axis and the point of launch; b is both the semiminor axis and the speed of launch.

The possibility of a connection with Newton's idea of transmuting the linear field into the gravitational one appears from the following surprising geometric fact. If we apply the mapping $z \rightsquigarrow z^2$ to an origin-centered ellipse, then the image is not some strange ugly shape, as one might expect, but rather another *perfect ellipse*; furthermore, this ellipse automatically has one focus at the *origin*. See FIG. 6. Before exploring the implications, let us verify this fact: squaring (3),

$$z \rightsquigarrow z^2 = (pe^{it} + qe^{-it})^2 = p^2e^{i2t} + q^2e^{-i2t} + 2pq.$$

The first two terms correspond to an origin-centered ellipse with foci at $\pm 2pq$; the last term therefore translates the left-hand focus to the origin.

Compare this purely geometric fact with Newton's dynamical argument. While leaving the orbit fixed, Newton moves the attracting point from the center to the focus, and deduces that the force law is transmuted from linear to inverse square; while leaving the attracting point fixed at the origin, $z \rightsquigarrow z^2$ transforms an orbit of the linear field into an orbit of the inverse square field. But we are only in a position to make the latter statement because of our prior knowledge of what the orbits in the two fields look like. *Is there instead some a priori reason why $z \rightsquigarrow z^2$ should map orbits of the linear field to orbits of the gravitational field?*

That there is indeed such a reason was apparently first discovered by K. Bohlin in 1911 ([5]); we shall therefore refer to this result as Bohlin's theorem. For a particularly clear account of the conventional explanation, see [4], p. 95. For more on the history of the idea, as well as applications to celestial mechanics, see [6], [7], and [8].

Granted that the origin-centered elliptical orbits for the linear field are easily derived (as above), Bohlin's theorem now allows us to view the geometric fate of these orbits under $z \rightsquigarrow z^2$ as a novel *explanation* of the elliptical motion of planets about the sun as focus. It is strange, though, that the only gravitational orbits we have managed to explain in this way are the ellipses; where are the hyperbolic orbits?

To resolve this, we must generalize Bohlin's result. In Section 6 we will show that gravitational orbits arise not only as the images of those in a linear field that is attractive, but also of those in a linear field that is *repulsive*. For the moment, though, let us simply use our prior knowledge of gravitational orbits (i.e. cheat) to empirically confirm that this result will indeed supply the missing hyperbolic orbits.

The differential equation for the repulsive linear field is $\ddot{z} = z$, the two basic solutions of which are $e^{\pm t}$. The (essentially) general solution can then be obtained by superposing conjugate amounts of these two:

$$z = \lambda e^t + \bar{\lambda} e^{-t}. \quad (4)$$

As illustrated on the left of Fig. 7, these are hyperbolic orbits with center (i.e. intersection of asymptotes) at the origin, and with foci at $\pm 2|\lambda|$. This may seem more familiar upon rewriting (4) as $z = a \cosh t + ib \sinh t$, where $a = \lambda + \bar{\lambda}$, and $ib = \lambda - \bar{\lambda}$. Both these numbers have the same physical significance as in the elliptical case: a = launch point; b = launch speed.

According to the asserted extension of Bohlin's result, we should seek the gravitational hyperbolas amongst the images under $z \rightsquigarrow z^2$ of the orbits in the

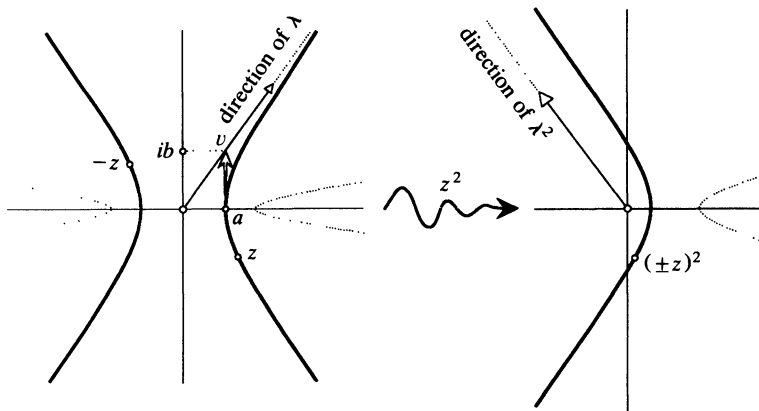


Figure 7.

repulsive linear field. What are these images?

$$z \rightsquigarrow z^2 = (\lambda e^t + \bar{\lambda} e^{-t})^2 = \lambda^2 e^{2t} + \bar{\lambda}^2 e^{-2t} + 2\lambda\bar{\lambda}. \quad (5)$$

The first two terms correspond to an origin-centered hyperbola with foci at $\pm 2|\lambda|^2$; the last term therefore translates the left-hand focus to the origin, apparently as hoped for.

We say ‘apparently’ because FIG. 7 presents a problem. While the solid hyperbola does indeed map to a gravitational orbit about the origin as attracting center, the dotted one does *not*. A possible interpretation of the dotted image would be that it represents a gravitational orbit about the *other* focus. This will not do, however, because we wish to think of the physical cause of the orbits as residing permanently at the origin, the other focus being devoid of physical influence. From this point of view it is clear that the dotted orbit must correspond to a *repulsive* field emanating from the origin. In the next section we will verify that this repulsive field is in fact inverse square, just like the attractive field that produces the solid orbit.

Next we shall use the conserved total energy E of the orbiting particle to characterize the distinction between those hyperbolas in the repulsive linear field that map to attractive orbits, and those that map to repulsive orbits.

If the particle’s speed is v , then since we shall always use a particle of unit mass, the kinetic energy contribution has the definite value $\frac{1}{2}v^2$, while the potential energy contribution is only defined up to a constant. The constant is fixed by arbitrarily assigning zero potential energy to some point in the plane. We shall make a very natural choice and assign zero energy to a point where the field vanishes: for a power law that diminishes with distance, at infinity; for a power law that increases with distance, at the origin.

In the repulsive linear field the total energy is then $E = (1/2)(v^2 - r^2)$, where $r = |z|$. Because E is constant, we may evaluate it at any point of the hyperbolic orbit. At launch, $v = b$ and $r = a$, so

$$E = \frac{1}{2}(b^2 - a^2).$$

But FIG. 7 informs us that the images that are attracted to the origin correspond to

values of λ for which λ^2 (the asymptotic direction of the image) points to the left. Since $4\lambda^2 = (a^2 - b^2) + i2ab$, we deduce that the image is attracted or repelled according as E is positive or negative. This characterization is equally applicable to the attractive linear field, for in that case $E = \frac{1}{2}(b^2 + a^2)$ is always positive, while the images are always attracted.

Returning to the repulsive field, we also anticipate that when $E = 0$ ($a^2 = b^2$), the image is neither attracted nor repelled, and is thus a straight line. Using (5), the reader may readily verify the truth of this conjecture.

Granted Bohlin's result and its extension, we have now explained the two principal kinds of motion in a gravitational field, and it only remains to explain the orbit which cannot decide if it's elliptical or hyperbolic: the androgynous parabola. Since ellipses and hyperbolas arise as images of orbits in linear force fields that are attractive and repulsive, respectively, we anticipate that the parabola will arise in the transitional case of zero force. The reader may now verify that the rectilinear orbits that arise in the absence of force do indeed map to parabolas with foci at the origin.

To end this section we will show that the effects of $z \rightsquigarrow z^2$ on both ellipses and hyperbolas are actually two equivalent facets of a single phenomenon. We shall argue that the result for ellipses implies the result for hyperbolas; the truth of the converse will then be apparent.

FIG. 8 shows an origin-centered hyperbola \mathcal{H} whose image under z^2 is sought. To find this image, draw the family of ellipses that is confocal to \mathcal{H} , and let p be an intersection of \mathcal{H} with one of these ellipses \mathcal{E} . Now a ray of light emitted from one of the foci towards p will be reflected directly towards the other focus by \mathcal{E} , and directly away from it by \mathcal{H} . But the reflected beam of light from a rotating mirror itself rotates twice as fast as the mirror, and thus to turn the beam through two right angles (as here) we must turn the mirror through one right angle. Thus \mathcal{H} is perpendicular to \mathcal{E} , and hence to the entire family of ellipses. Conversely, it is clear that any orthogonal trajectory through the ellipses must be a confocal hyperbola. Now apply $z \rightsquigarrow z^2$. By the assumed result, the family of ellipses is mapped to another confocal family of ellipses, one focus being at the origin. But because $z \rightsquigarrow z^2$ is *conformal*, the image of \mathcal{H} must be an orthogonal trajectory through the new family of ellipses. Q.E.D.

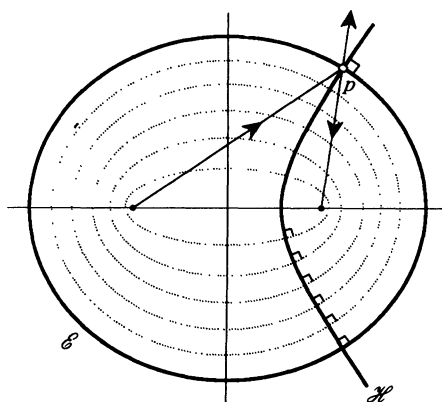


Figure 8.

4. NEWTON'S GEOMETRY OF FORCE. The second fundamental physical constant associated with a central orbit is its angular momentum h . If r is the distance of the particle from the center of force, and ω its angular speed about this center, then $h = r^2\omega$. The geometric interpretation of this quantity is that it's simply twice the "areal speed" \mathcal{A} = (the rate at which the radius vector sweeps out area).

There is a complication associated with Bohlin's theorem that arises from the following basic theorem of Newton [Props. I & II]: \mathcal{A} , and therefore h , will remain constant if and only if the force field is central. Suppose that we watch a film of a particle orbiting in the linear field along the left-hand ellipse of FIG. 6, while its image travels round the right-hand ellipse. Although the complete path traced by the image is indeed a genuine gravitational orbit, the way in which the image moves in *time* is physically impossible! For, letting tilde indicate a quantity associated with the image under $z \rightsquigarrow z^2$, we find that

$$\tilde{\mathcal{A}} = \frac{1}{2}\tilde{r}^2\tilde{\omega} = \frac{1}{2}(r^2)^2(2\omega) = 2r^2\mathcal{A},$$

thereby making it impossible for both \mathcal{A} and $\tilde{\mathcal{A}}$ to remain constant. If \mathcal{A} is held constant [physical motion for preimage] while we make our film, then when we play it back, the projector must be continually speeded up and slowed down (in proportion to r^2) to make the image appear to sweep out area at a constant rate.

In light of this fact it is curious that the conventional explanation of Bohlin's result characterizes a force field by means of a *temporal* differential equation. To overcome the above difficulty, it becomes necessary to introduce a fictitious time coordinate into the image orbit (cf. our film projector trick) in such a way that area is swept out at a constant rate with respect to it. It therefore seems natural to seek an alternative explanation that avoids explicit mention of time, and instead directly addresses the geometry of the orbits. Before embarking on the details of this new explanation, let us outline the strategy that will be followed.

In the absence of force, a particle will move in a straight line; *bending* is therefore the manifestation of force, and this can be quantified in terms of the curvature of the orbit. In this section we derive the formula that relates the force to the curvature, and give three interesting examples of its use. In the next section we discover how the curvature of an orbit is transformed under an analytic mapping. Finally, having understood both the transformation of curvature and its relationship to force, in Section 6 we deduce the transmutation of force that is effected by an analytic mapping.

The relationship between force and curvature could be derived rapidly by appealing to standard elementary results in dynamics. However, continuing in the spirit of the introduction, we choose instead to illustrate the *Principia's* method by deriving it from scratch, using little more than similar triangles.

In FIG. 9(A), which shows a particle orbiting in a central force field emanating from C , we have drawn QS and CY parallel to the normal at P , and QR parallel to the radius CP . Our problem is to determine from the shape of the orbit the force F that is acting on the particle when it is at the point P , and ultimately to express the result in terms of the curvature κ . Initially, we shall simply follow Newton.

After a short time Δt the particle will have moved along the orbit from P to Q , and the radius will have swept out an area $(1/2)h \Delta t$. The net motion from P to Q can be thought of as compounded of the force-free motion PR along the tangent, together with the force-induced deflection QR in the direction of the force acting at P . This view of events becomes increasingly accurate as we shorten Δt , for as Q

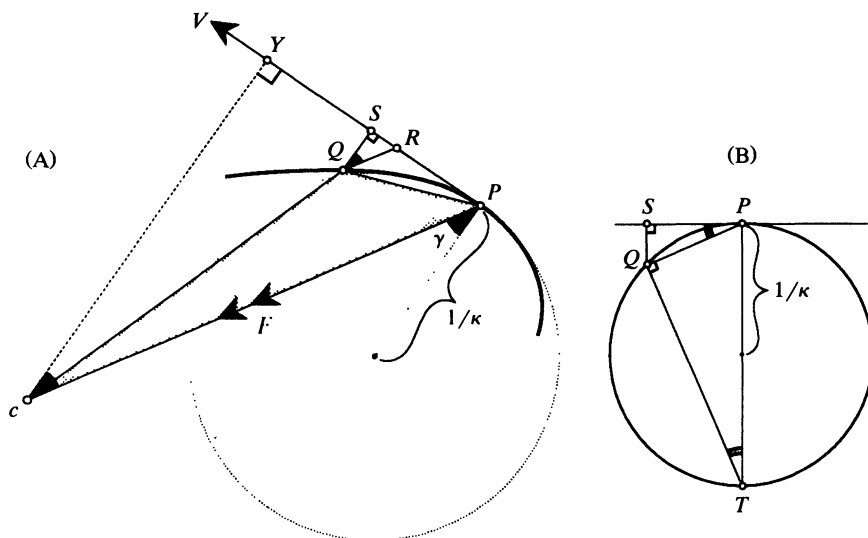


Figure 9.

is taken closer and closer to P , the force (including its direction) acting in the arc PQ becomes more and more nearly constant, and coincides with that at P .

Before pursuing this idea, and in the interests of brevity, let us adopt the following notation and form of words as being equivalent to the statement

$$\text{“} \lim_{Q \rightarrow P} \left(\frac{X}{Y} \right) = 1 \text{”}:$$

“ $X \asymp Y$ ” \Leftrightarrow “ X and Y become equal as Q coalesces with P ”.

The basic theorems on limits inform us that “ \asymp ” inherits many of the properties of “ $=$ ”. This simple notational device merely paraphrases the idea which Newton expresses verbally in the Scholia on pages 35 and 37 of the *Principia*. It allows us to draw on the intuitive power of his infinitesimal geometry while continuing to pay lip-service to the tyrannical ϵ - δ legacy of Cauchy and Weierstrass.

Returning to the investigation, and recalling that under the action of a constant force, *distance* = (half) \cdot (acceleration) \cdot (time squared), we deduce that the force at P is given by

$$F \asymp \left[\frac{2 \cdot QR}{(\Delta t)^2} \right]. \quad (6)$$

We may now render this formula purely geometrical by using the area law to eliminate time:

$$\begin{aligned} \frac{1}{2} h \Delta t &= (\text{area of sector } CPQ) \\ &\asymp (\text{area of shaded triangle } CPQ) \\ &\asymp \frac{1}{2} (PQ \cdot CY), \end{aligned}$$

where the last ‘equality’ follows from observing the evolution as $Q \rightsquigarrow P$ of the altitude from C to the base PQ . Substituting for Δt in (6), we obtain one of the formulae [Cor. II., Prop. VI] that lie at the heart of Newton’s geometric investigation of force in the *Principia*:

$$F \asymp h^2 \left[\frac{2 \cdot QR}{PQ^2 \cdot CY^2} \right].$$

In order to re-express Newton's formula in terms of curvature, consider FIG. 9(B). From the similar triangles PSQ and TQP , we find that

$$\frac{PQ}{PT} = \frac{QS}{PQ} \Rightarrow \kappa = \left(\frac{2 \cdot QS}{PQ^2} \right).$$

This formula may also be applied to the non-circular orbit in FIG. 9(A), for as Q coalesces with P , it will yield the curvature of the orbit at P . Next, observe that we may re-express QR as $(QS \cdot \sec \gamma)$, and CY as $(CP \cdot \cos \gamma)$, where γ is the angle between the normal and the radius at P . Thus

$$F \asymp h^2 \left(\frac{2 \cdot QS}{PQ^2} \right) \left(\frac{\sec^3 \gamma}{CP^2} \right),$$

and writing r for CP , we arrive at our destination [cf. Prop. VII]:

$$F = h^2 \left(\frac{\kappa \sec^3 \gamma}{r^2} \right). \quad (7)$$

In this formula, γ is to be taken as acute so that $\sec \gamma$ is always positive; in the next section we shall introduce a sign for κ so as to distinguish between attractive and repulsive forces.

Before describing some applications of this formula, let us note another result that we shall need later. If v is the speed of the particle at P , then [Cor. I, Prop. I]

$$v \asymp \left(\frac{PQ}{\Delta t} \right) \asymp \left(\frac{h}{CY} \right),$$

and therefore

$$v = h \left(\frac{\sec \gamma}{r} \right). \quad (8)$$

As our first application of (7), let us ask and answer the question, what force field can yield a circular orbit that passes through the center of attraction? See FIG. 10(A). Since $\sec \gamma = (2\rho/r)$ and $\kappa = (1/\rho)$, we immediately find [Cor. I,

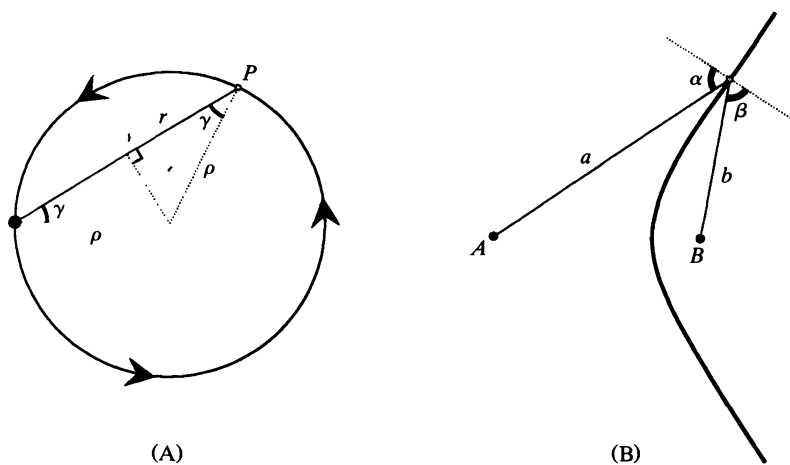


Figure 10.

$$F = \left(\frac{8h^2\rho^2}{r^5} \right).$$

The significance of the $(1/r^5)$ field, and this particular orbit within it, will be discussed in Section 6.

For our second example, turn to Fig. 10(B). Granted that such a hyperbola is a gravitational orbit about the focus B as attracting center, you may now verify (using $\alpha = \beta$) our previous claim that if the orbit is instead due to a repulsive field emanating from the other focus A , then this field must also be inverse square.

As the last of our three examples, let us verify Newton’s “transmutation formula” (1). Reconsider FIG. 4. Let fall the perpendicular AR from A onto the tangent PRQ , and let α and β be the angles PAR and QAR respectively. Then

$$AP \cdot \cos \alpha = AR = AQ \cdot \cos \beta \Rightarrow \frac{\sec \beta}{\sec \alpha} = \frac{AQ}{AP}.$$

Using (7) to write down the ratio of the new force to the old one, the result (1) is now readily verified.

5. ANALYTIC TRANSFORMATION OF CURVATURE. Suppose that an analytic mapping $f(z)$ acts on an orbit of curvature κ , yielding an image orbit of curvature $\tilde{\kappa}$. What is $\tilde{\kappa}$ in terms of κ and $f(z)$? This is the question that we must answer if we are to understand the relationship between the force fields that hold the preimage and image in their respective orbits. While we could certainly answer this question by calculation, we choose instead to present a novel solution that is entirely geometric.

The means by which the curvature of an orbit will be determined is illustrated in FIG. 11(A), where a particle is seen orbiting in a counterclockwise circle. As shown, ξ and ζ are two successive chords of equal length in the direction of motion, and ϵ is the angle of turning from ξ to ζ . Because the angle ϕ that each of these chords subtends at the center must equal the turning angle ϵ ,

$$\frac{1}{\kappa} \cdot \epsilon = (\text{the arc } PQ) \asymp |\xi|.$$

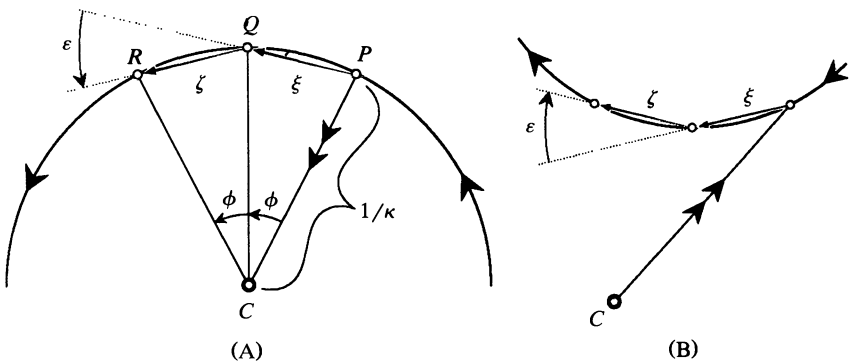


Figure 11.

The curvature can therefore be expressed as

$$\kappa \asymp \frac{\epsilon}{|\xi|}. \quad (9)$$

It is clear that this construction can immediately be applied to a general non-circular orbit: as the equal lengths of ξ and ζ diminish, and as Q and R coalesce with P , we obtain the curvature of the orbit at P .

Before continuing, let us introduce some interrelated conventions in order to systematically distinguish between attractive and repulsive forces. Although time is essentially to be suppressed in what follows, it is still helpful to picture the particle as moving along the orbit, and our first convention is to insist that the sense of this motion be *counterclockwise* about the center of force.

With this convention in place, (9) now attributes a definite sign to κ . For example, if the center of force is at C in both FIG. 11(A) and FIG. 11(B), then κ is positive in the first, and negative in the second; we also notice that the responsible forces are attractive and repulsive, respectively. Thus, by using this signed curvature in (7), we may identify a positive F with attraction, and a negative F with repulsion.

We now return to the original problem, which is both depicted and solved in FIG. 12. The left-hand figure shows the preimage orbit of curvature κ , and the top figure shows its image under $f(z)$. Just as ξ connects P and Q , so $\tilde{\xi}$ connects the image points \tilde{P} and \tilde{Q} , and we shall say that $\tilde{\xi}$ is the “image” of ξ . Likewise, just

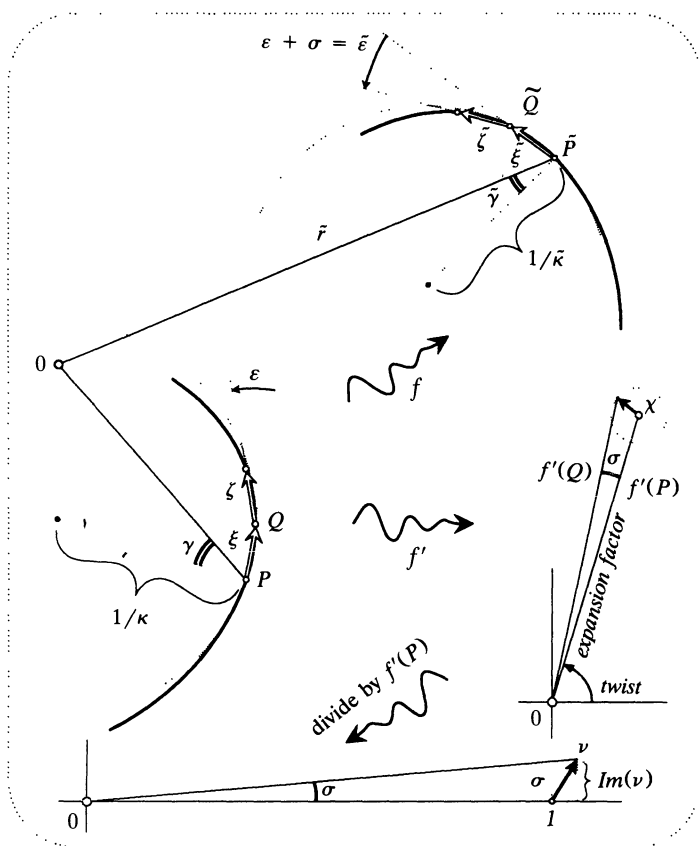


Figure 12.

as ξ and ζ yield the curvature of the preimage, so their images yield $\tilde{\kappa}$:

$$\tilde{\kappa} \asymp \frac{\tilde{\epsilon}}{|\tilde{\xi}|}, \quad (10)$$

and our problem therefore reduces to finding $\tilde{\epsilon}$ and $|\tilde{\xi}|$.

The distinguishing characteristic of an analytic function is that its local effect is simply to expand and twist. Thus if we imagine a microscopic circle centered at P , its image under such a mapping will not be a tiny ellipse (as it would be in general), rather it will be another tiny *circle* centered at \tilde{P} . As illustrated in the right-hand figure, the information of the expansion factor and the twist that carry the former circle into the latter is stored in the single complex number $f'(P) = (\text{expansion factor})e^{i(\text{twist})}$. Part of the problem is therefore easily solved:

$$|\tilde{\xi}| \asymp (\text{expansion factor}) \cdot |\xi| = |f'(P)| \cdot |\xi|. \quad (11)$$

The more interesting and difficult part of the problem is to find $\tilde{\epsilon}$. If ξ and ζ both underwent precisely the same twist, then the turning angle $\tilde{\epsilon}$ for the images would equal the original angle ϵ . However, the twist at Q will differ very slightly, say by σ , from that at P . Thus

$$\tilde{\epsilon} = \epsilon + (\text{extra twist}) = \epsilon + \sigma. \quad (12)$$

This angle σ (which we must find) is illustrated in the right-hand figure.

To find σ , we appeal to the astonishing fact that if a mapping sends tiny circles to tiny circles, then *so does its derivative*: in more conventional language, an analytic mapping is infinitely differentiable. Thus $f'(z)$ maps a tiny circle centered at P to a tiny circle [dotted] centered at $f'(P)$, and the expansion and twist that carries the former to the latter is encoded as $f''(P)$: if (as shown) χ is the image of ξ under $f'(z)$, then

$$\chi \asymp f''(P) \cdot \xi.$$

Knowing χ , we are now very close to finding the extra twist σ , for we observe that it is the angle at the origin in the triangle of the right-hand figure, the sides of which are the known quantities $f'(P)$ and χ . It is easier to obtain an expression for σ if we first rotate this triangle to the real axis. This rotation is achieved quite naturally (see the bottom figure) by dividing by $f'(P)$; the sides of the triangle now become 1 and $\nu = [\chi/f'(P)]$. Because σ equals the almost vertical arc through 1, we see from the figure that

$$\sigma = \text{arc} \asymp \text{Im}(\nu) = \text{Im}\left[\frac{\chi}{f'(P)}\right] \asymp \text{Im}\left[\frac{f''(P) \cdot \xi}{f'(P)}\right].$$

Thus, from (10), (11), and (12), and taking evaluation at P as understood, we obtain

$$\tilde{\kappa} \asymp \frac{\left(\text{Im}\left[\frac{f'' \cdot \xi}{f'}\right] + \epsilon\right)}{|f'| \cdot |\xi|}.$$

Finally, using (9), and writing $\hat{\xi}$ for the unit tangent at P , we arrive at our transformation formula:

$$\boxed{\tilde{\kappa} = \frac{1}{|f'|} \left(\text{Im}\left[\frac{f'' \cdot \hat{\xi}}{f'}\right] + \kappa \right)}. \quad (13)$$

While I doubt that this interesting formula can be new, I confess that I have been unable to find it elsewhere.

Let us immediately do an example. With $f(z) = e^z$, and writing $z = x + iy$, we find that

$$\tilde{\kappa} = e^{-x}(\sin \phi + \kappa),$$

where ϕ is the angle that the tangent makes with the horizontal. FIG. 13 shows three line-segments on the left, and their images under $z \rightsquigarrow e^z$ on the right. The reader is encouraged to verify the accord between this figure and our formula.

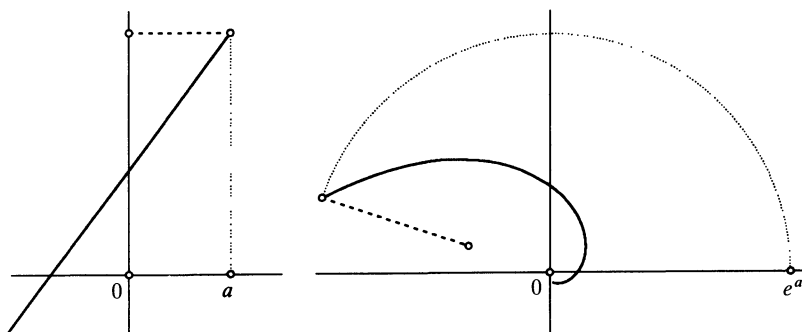


Figure 13.

In order to understand Bohlin's theorem and its generalizations, we need to know the transformation law for the power mappings $f(z) = z^m$. Substituting this into (13), we obtain

$$\tilde{\kappa} = \frac{1}{|m|r^{m-1}} \left[(m-1) \operatorname{Im} \left(\frac{\hat{\xi}}{z} \right) + \kappa \right] = \frac{1}{|m|r^{m-1}} \left[(m-1) \frac{\sin \psi}{r} + \kappa \right],$$

where ψ is the angle from z to the directed unit tangent $\hat{\xi}$.

With the origin as center of force, our convention of counterclockwise motion is now equivalent to $0 < \psi < \pi$, and so $\sin \psi$ is positive. Using this fact, we may rewrite the above result in its final form:

$$\tilde{\kappa} = \frac{1}{mr^{m-1}} \left[(m-1) \frac{\cos \gamma}{r} + \kappa \right], \quad (14)$$

where γ is the acute angle, familiar from the force formula (7), between the normal and the radius.

Notice that in the denominator of this formula we have written m instead of $|m|$, which would be the strict deduction from (13). Thus the formula agrees with the strict deduction when m is positive, and is opposite to it when m is negative. The appropriateness of this choice stems from the convention of counterclockwise motion. With $m > 0$, conventional motion for the preimage results in conventional motion for the image, as depicted in FIG. 12. However, with $m < 0$, the sense of the motion is *reversed* by the mapping, and the value of $\tilde{\kappa}$ furnished by (13) is therefore opposite to that required by our convention.

As a simple example of this, consider the case of inversion, $z \rightsquigarrow (1/z)$. A vertical line is mapped to a circle through the origin having *positive* curvature

according to our convention. To check (14) we simply put $m = -1$ and $\kappa = 0$:

$$\tilde{\kappa} = +2r \cos \gamma.$$

Thus the formula yields the correct sign, and a picture quickly reveals that the actual value is also correct.

6. THE TRANSMUTATION OF FORCE. At last we are in a position to find out the effect of an analytic mapping on force itself. Applying $z \rightsquigarrow f(z)$,

$$F = h^2 \left(\frac{\kappa \sec^3 \gamma}{r^2} \right) \rightsquigarrow \tilde{F} = \tilde{h}^2 \left(\frac{\tilde{\kappa} \sec^3 \tilde{\gamma}}{\tilde{r}^2} \right),$$

where $\tilde{\kappa}$ is known from (13), and where $\tilde{r} = |f(P)|$ is the distance to the image, as illustrated in FIG. 12. Note that a general conformal mapping will *not* preserve the angle γ . However, if we restrict ourselves to the *power mappings* $f(z) = z^m$, then the ray through P is mapped to the ray through \tilde{P} , and therefore $\tilde{\gamma} = \gamma$.

As our first example of the above idea, let us consider $f(z) = z^2$, for this should lead us to Bohlin's result. Putting $m = 2$ in (14), and then substituting for the original force and speed from (7) and (8), we deduce that the force responsible for the image orbit is

$$\tilde{F} = \tilde{h}^2 \cdot \frac{\frac{1}{2} \left[\frac{\cos \gamma}{r^2} + \frac{\kappa}{r} \right] \sec^3 \gamma}{\tilde{r}^2} = \left(\frac{\tilde{h}}{h} \right)^2 \frac{\left[\frac{1}{2} v^2 + \frac{1}{2} r F \right]}{\tilde{r}^2}.$$

Even if F is a simple power law, generally this \tilde{F} will not be. However, if and *only* if the original force field is the attractive or repulsive *linear* one ($F = \pm r$), the numerator in the above expression magically becomes the constant total energy E of the particle in the original field:

$$\tilde{F} = \left(\frac{\tilde{h}}{h} \right)^2 \frac{E}{\tilde{r}^2} !$$

The image, therefore, moves in a field that is *inverse-square*. Furthermore, if the original orbit has positive energy then its image is attracted to 0, while if it has negative energy then its image is repelled by 0. We have therefore successfully explained all of our empirical findings in Section 3.

Before describing the general result (due to Arnol'd, [3]) on "dual" pairs of power laws, we shall treat one further special case that arises from the following interesting question. Might there exist a force field that is *self-dual* in the sense that the images of orbits within it (under a complex power mapping) would simply be new orbits in the *same* field?

If this were possible then clearly the responsible mapping would have to be *self-inverse*. Discarding the identity mapping as trivial, the sole possibility is therefore inversion: $z \rightsquigarrow (1/z)$. Putting $m = -1$ into (14), and noting that $\tilde{r} = (1/r)$, we find that

$$\tilde{F} = +4H^2 \frac{\left[\frac{1}{2} v^2 - \frac{1}{4} r F \right]}{\tilde{r}^5},$$

where H stands for the ratio of the angular momenta, (\tilde{h}/h) . In order for the numerator to become the total energy, the potential energy must be $-\frac{1}{4} r F$, and therefore the original force field must be the attractive or repulsive inverse fifth power: $F = \pm(1/r^5)$. As anticipated, the force acting on the image is then *also*

inverse-fifth, and furthermore it is attractive or repulsive according as the original energy is positive or negative.

We previously found an orbit of this field in FIG. 10(A); its very special character can be seen from the fact that under inversion it maps to a force-free straight line. For pictures of more general orbits, as well as their mathematical classification, the interested reader may consult [10].

We turn now to the general case. Suppose that the mapping $z \rightsquigarrow z^m$ acts on orbits in a force field F . Recalling that $\tilde{r} = r^m$, we find that

$$\tilde{F} = \frac{2(m-1)}{m} H^2 \left[\frac{1}{2} v^2 + \frac{rF}{(2m-2)} \right] \tilde{r}^{[(2/m)-3]}. \quad (15)$$

In order to obtain the total energy in this expression, the potential energy must be $[rF/(2m-2)]$. Only for a power law is this energy proportional to rF , and, in greater detail, for $F = \pm r^A$ it equals $[rF/(A+1)]$. We deduce that the mapping that effects the transmutation must be related to the original force field by $(2m-2) = (A+1)$. If \tilde{A} stands for the exponent in the image power law, so that $\tilde{A} = [(2/m)-3]$, we may then summarize our findings as follows.

Associated with each power law (exponent A) there is precisely one power law (exponent \tilde{A}) that is “dual” in the sense that orbits of the former are mapped to orbits of the latter by $z \rightsquigarrow z^m$, and the relationships between the forces and the mapping are:

$$(A+3)(\tilde{A}+3) = 4 \quad \text{and} \quad m = \frac{(A+3)}{2}. \quad (16)$$

Observing the coefficient $[(m-1)/m]$ in (15), we further conclude that (in general) positive energy orbits in either the attractive or repulsive field $F = \pm r^A$ map to attractive orbits in the dual field, while negative energy orbits map to repulsive ones. However, if $-3 < A < -1$ (e.g. gravity) then these roles are reversed. In all cases, zero energy orbits map to force-free rectilinear orbits.

Letting $[A, \tilde{A}]$ stand for a pair of dual force laws, we see from (16) that amongst *integer* exponents (and excluding $z \rightsquigarrow z$ as trivial) there are only three such pairs:

$$[1, -2] \quad [-4, -7] \quad [-5, -5]. \quad (17)$$

Of these only one is new to us, namely, $[-4, -7]$; the mapping in this case is $z \rightsquigarrow (1/\sqrt{z})$.

7. CONCLUDING REMARKS. (I) In his book, Arnol’d alludes to, but does not state, a connection between (16) and a result of Newton’s [Prop. XLV]. While I cannot be sure of what *Arnol’d* had in mind, let me at least point out a connection.

Now it so happened that the solar system formed in such a way that the planetary and lunar orbits are almost circular. For this reason Newton undertook the investigation of almost circular orbits for a general power law $F = r^A$. He found that the angle Φ between successive turning points of r (aphelion to perihilion, or vice versa) is given by the formula

$$\Phi = \frac{\pi}{\sqrt{A+3}}.$$

With $A = 1$ or -2 , this value of Φ is exact (as you may verify in FIG. 6) even for highly non-circular orbits.

Next we observe that turning points may be identified by the property $\gamma = 0$. If $z \rightsquigarrow z^m$ then $\tilde{\gamma} = \gamma$, and so a turning point is mapped to a turning point, and no new turning points are created in the process. We deduce that the angular separation of turning points on the image curve is $\tilde{\Phi} = m\Phi$. But if the existence of duality is assumed [of course this is only possible in hindsight] then this image curve will be an orbit for some new power law, say with exponent \tilde{A} , and therefore

$$\tilde{\Phi} = \frac{\pi}{\sqrt{\tilde{A} + 3}}.$$

Equating these two expressions for $\tilde{\Phi}$, we recover half of the information contained in (16) but expressed in a different form:

$$m^2 = \left(\frac{A + 3}{\tilde{A} + 3} \right).$$

(II) With the exception of the pair $[1, -2]$, the integer power laws ($A = -4, -5, -7$) that are singled out in (17) are rather mysterious. Their physical significance is unknown to me, and that in itself is strange, for the music of mathematics is seldom played without an accompanying echo being heard in Nature. At the purely mathematical level, however, there is clearly more here than first meets the eye. For example, consider the following question (see [9]): for which integer power laws that diminish with distance is the polar equation of the orbit expressible in terms of elliptic functions? It turns out that there are precisely three. They are: -4 , -5 , and -7 !

ACKNOWLEDGMENTS. First I should like to thank Dr. Stanley Nel for his enthusiastic reception of the main idea, and for his helpful comments on the first draft. I also thank Dr. Subrahmanyam Chandrasekhar for helpful observations on the Newtonian portion of the paper. Lastly, I thank Dr. James Finch for his patient and very able assistance in overcoming various problems with T_EX, and in particular for his figuring out how to import my diagrams, which I had created using “CorelDRAW”.

REFERENCES

1. *Newton's Principia*, U. of California Press, 1934.
2. R. Westfall, *Never at Rest*, Cambridge U. Press, 1980.
3. V. I. Arnol'd and V. A. Vasil'ev, Newton's Principia read 300 years later, *Notices Amer. Math. Soc.*, 36 (1989) 1148–1154.
4. V. I. Arnol'd, *Huygens & Barrow, Newton & Hooke*, Birkhäuser Verlag, 1990.
5. K. Bohlin, Note sur le problème des deux corps et sur une intégration nouvelle dans le problème des trois corps, *Bull. Astr.*, 28 (1911) 113–119.
6. K. R. Meyer, The Geometry of Harmonic Oscillators, this *Monthly*, 97 (1990) 447–465.
7. D. G. Saari, A visit to the Newtonian N -body problem via elementary complex variables, this *Monthly* 97 (1990) 105–119.
8. V. Szebehely, *Theory of Orbits*, Acad. Press, New York, 1967. Chapter 3.
9. H. Goldstein, *Classical Mechanics*, second Ed., Addison-Wesley, 1980. Chapter 3.
10. W. D. MacMillan, *Statics and the Dynamics of a Particle*, McGraw-Hill, 1927. Chapter XII, Section V. Also by the same author, see *Amer. J. Math.*, Vol. XXX, 282–306.

Mathematics Department
University of San Francisco
San Francisco, CA 94117-1080

Note added in proof: We attributed the general duality law (16) to Arnol'd, but it was in fact discovered by Edward Kasner in or before 1909. It appears in his *Differential-Geometric Aspects of Dynamics*, published by the A.M.S. in 1913.

A Note on Diophantine Representations

Christoph Baxa

In 1900 David Hilbert asked for an algorithm to decide whether a given diophantine equation is solvable or not and put this problem tenth in his famous list of 23. In 1970 it was proved that such an algorithm cannot exist, i.e. the problem is recursively undecidable. Proof was supplied by Yu. V. Matijasevič [11], heavily leaning on results arrived at by M. Davis, J. Robinson and H. Putnam [18], [4]. This was accomplished by proving that any recursively enumerable set $A \subseteq \mathbb{N} = \{0, 1, 2, \dots\}$ can be represented in the following form: There exists a polynomial $p(x, x_1, \dots, x_n)$ with $n \geq 0$ such that $a \in A$ if and only if $p(a, x_1, \dots, x_n) = 0$ is solvable for particular nonnegative integers x_1, \dots, x_n , i.e.

$$a \in A \Leftrightarrow \exists x_1, \dots, x_n \geq 0: p(a, x_1, \dots, x_n) = 0.$$

Therefore, the set A equals the set of parameters for which the equation $p = 0$ is solvable. Employing an idea of H. Putnam [16] this can be reformulated as follows. Put $q(x, x_1, \dots, x_n) = x(1 - p(x, x_1, \dots, x_n)^2)$, then A equals the set of positive values of q , where its variables range over the nonnegative integers. Among the recursively enumerable sets are many for which such representation is surprising. I will name some examples which are of importance in number theory.

- (1) The primes and their recursively enumerable subsets, most outstanding Fermat-, Mersenne- and twin-primes.
- (2) The set of partial denominators of the continued fraction expansion of numbers as e , π and $\sqrt[3]{2}$. (Whereas for e this is known to equal $\{1\} \cup \{2, 4, 6, \dots\}$, there is only computer-based research regarding the other numbers, see e.g. [5], [9] and the references therein.)
- (3) The set of all natural numbers $n \geq 3$ for which $x^n + y^n = z^n$ is solvable in $\mathbb{N} \setminus \{0\}$.

Of these examples, only those mentioned in the first paragraph have received extensive treatment (cf. [12], [8], [13], [6]). Furthermore, there exists a diophantine equation whose unsolvability is equivalent to the Riemann hypothesis (cf. [3], [14]). It is the purpose of this note to present some interesting corollaries of theorems proved in [8], dealing with the set of twin-primes and Fermat's last theorem. The interested reader, who is looking for background-information on the solution of Hilbert's tenth problem, is referred to [2], [3], [19] and the sixth chapter of [10].

Constructing a diophantine representation of Fermat's last theorem does not pose a serious problem, given our present knowledge about diophantine representations. For example, it is not difficult to prove the following.

Theorem 1. *Fermat's last theorem holds if and only if the following system of diophantine equations is unsolvable, i.e. there are no nonnegative integers*

$a, c, d, e, f, g, h, i, j, l, n, o, p, r, u, v, w, x, y, z$ such that

- (1) $v^2 - (a^2 - 1)w^2 = 1$,
- (2) $u^2 = 16(a^2 - 1)r^2w^4 + 1$,
- (3) $e^3(e + 2)(a + 1)^2 + 1 = o^2$,
- (4) $w = n + p + 3$,
- (5) $f + g = h$,
- (6) $e = x + y + z + n + f + g + h + 5$,
- (7) $(v + cu)^2 = ((a + u^2(u^2 - a))^2 - 1)(n + 4dw + 3)^2 + 1$,
- (8) $v = f + w(a - x - 1) + i(2a(x + 1) - (x + 1)^2 - 1)$,
- (9) $v = g + w(a - y - 1) + j(2a(y + 1) - (y + 1)^2 - 1)$,
- (10) $v = h + w(a - z - 1) + l(2a(z + 1) - (z + 1)^2 - 1)$.

Proof: Fermat's last theorem is false if and only if there exist $n, x, y, z \geq 0$ such that $(x + 1)^{n+3} + (y + 1)^{n+3} = (z + 1)^{n+3}$. This holds if and only if there exist $f, g, h, n, x, y, z \geq 0$ such that $f = (x + 1)^{n+3}$, $g = (y + 1)^{n+3}$, $h = (z + 1)^{n+3}$ and $f + g = h$. Whereas the last equation is just equation (5) the three exponentiations are encoded in equations (1)–(4) and (6)–(10). This part of the proof can be accomplished by using a standard argument and is left to the reader. Essentially, Corollary 2.6 and the Lemmata 2.3 and 2.4 in [8] contain all the information needed. Parts of the proof of Theorem 2.12 in [8] can serve as a model.

The next theorem describes a very economic representation of the factorial.

Theorem 2. *For any positive integers g and k in order that $g = k!$ it is necessary and sufficient that there exist nonnegative integers $a, b, c, d, e, f, h, i, j, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z$ such that*

- (1) $q = wz + h + j$,
- (2) $z = g(h + j) + h$,
- (3) $e = p + q + z + 2n$,
- (4) $x^2 = (a^2 - 1)y^2 + 1$,
- (5) $m^2 = (a^2 - 1)l^2 + 1$,
- (6) $l = k + i(a - 1)$,
- (7) $n + l + v = y$,
- (8) $(2k)^3(2k + 2)(n + 1)^2 + 1 = f^2$,
- (9) $e^3(e + 2)(a + 1)^2 + 1 = o^2$,
- (10) $u^2 = 16(a^2 - 1)r^2y^4 + 1$,
- (11) $(x + cu)^2 = ((a + u^2(u^2 - a))^2 - 1)(n + 4dy)^2 + 1$,
- (12) $m = p + l(a - n - 1) + b(2a(n + 1) - (n + 1)^2 - 1)$,
- (13) $x = q + y(a - p - 1) + s(2a(p + 1) - (p + 1)^2 - 1)$,
- (14) $pm = z + pl(a - p) + t(2ap - p^2 - 1)$.

Proof: This is an expanded version of Lemma 2.11 in [8]. Equations (i), (ii) and (iii) are the same as equations (1), (2) and (8), save changing f to g and replacing \square by f^2 . The equations (3)–(7) and (9)–(14) are used to encode the three exponentiations (iv), (v) and (vi). As before, this can be accomplished using Corollary 2.6 and the Lemmata 2.3 and 2.4 in [8], again using the proof of Theorem 2.12 in [8] as a model.

The equations of Theorem 2 bear a very strong resemblance with those of Theorem 2.12 in [8], in fact they are identical up to numbering and equations (2). The authors of [8] made use of Wilson's theorem, i.e. the fact that $k + 1$ is a prime if and only if $k! \equiv -1 \pmod{k + 1}$. This can be restated as follows:

$k + 1$ is a prime number if and only if there exists $g \geq 0$ such that $k! + 1 = (g + 1)(k + 1) = gk + g + k + 1$, i.e. $k! = gk + g + k$. Utilizing the diophantine representation of $k!$ as stated in Theorem 2 in this version of Wilson's theorem results in Theorem 2.12 and thus in changing the appearance of equation (2) as stated above.

Obviously, this enables us to construct diophantine representations for (recursively enumerable) sets of primes just by adding their defining properties in the form of equations. The first sets one might think of are the Fermat- and the Mersenne-primes. Clearly diophantine representations of the equations $k + 1 = 2^n + 1$ and $k + 1 = 2^n - 1$ could be added, but for these sets special primality-tests are known. Employing these J. P. Jones [6] constructed the following short polynomials.

Theorem 3. *The set of Mersenne-primes equals the set of positive values of the polynomial p_M , where its variables range over the non-negative integers. Likewise, the same holds for the Fermat-primes and the polynomial p_F .*

$$\begin{aligned}
 p_M(a, b, c, d, f, g, h, i, j, k, l, m, n) \\
 = n \Big(1 - (4b + 3 - n)^2 - b \Big((2 + hn^2 - a)^2 \\
 + (n^3 d^3 (nd + 2)(h + 1)^2 + 1 - m^2)^2 \\
 + (db + d + chn^2 + g(4a - 5) - kn)^2 \\
 + ((a^2 - 1)c^2 + 1 - k^2 n^2)^2 + (4(a^2 - 1)i^2 c^4 + 1 - f^2)^2 \\
 + \left((kn + lf)^2 - \left((a + f^2(f^2 - a))^2 - 1 \right) (b + 1 + 2jc)^2 - 1 \right)^2 \Big) \Big),
 \end{aligned}$$

$$\begin{aligned}
 p_F(a, b, c, d, e, f, g, h, i, j, k, l, m, n) \\
 = (6g + 5) \Big(1 - (bh + (a - 12)c + n(24a - 145) - d)^2 \\
 - (16b^3 h^3 (bh + 1)(a + 1)^2 + 1 - m^2)^2 \\
 - (3g + 2 - b)^2 - (2be + e - bh - 1)^2 \\
 - (k + b - c)^2 - ((a^2 - 1)c^2 + 1 - d^2)^2 \\
 - (4(a^2 - 1)i^2 c^4 + 1 - f^2)^2 \\
 - \left((d + lf)^2 - \left((a + f^2(f^2 - a))^2 - 1 \right) (b + 2jc)^2 - 1 \right)^2 \Big).
 \end{aligned}$$

All even perfect numbers are of the form $2^{p-1}(2^p - 1)$, where $2^p - 1$ is a Mersenne-prime, so this diophantine representation of Mersenne-primes may be used to construct one for even perfect numbers. The structure of p_M indicates that $n = 4b + 3$ for all positive values of p_M . This implies $4b + 3 = 2^p - 1$ and $2b + 2 = 2^{p-1}$ for some prime number p . So a like theorem [6] holds for the set of even perfect numbers and the polynomial $(2b + 2)p_M(a, b, c, d, f, g, h, i, j, k, l, m, n)$.

Now I want to say a few words about the set of twin-primes. The easiest way of representation would be by duplicating all equations for primes to the effect that $p, p + 2$ are twin-primes if and only if p and $p + 2$ are primes. A less clumsy way of dealing with the problem is using the following theorem.

Theorem 4. Let $n \geq 2$. The integers $n, n + 2$ form a pair of twin-primes if and only if

$$4((n - 1)! + 1) + n \equiv 0 \pmod{n(n + 2)}.$$

This is just a variant of Wilson's theorem. It was originally proved by P. A. Clement [1]; a more elaborate proof can be found in P. Ribenboim's book [17]. Let us now call a number p a "younger twin-prime" if $p, p + 2$ is a pair of twin-primes.

Corollary 5. The set of younger twin primes is identical with the positive values assumed by the polynomial p_T , where a, \dots, z range over the nonnegative integers.

$$\begin{aligned} p_T(a, \dots, z) = & (k + 2) \left(1 - (wz + h + j - q)^2 - ((g + 1)(h + j) + h - z)^2 \right. \\ & - (p + q + z + 2n - e)^2 - (e^3(e + 2)(a + 1)^2 + 1 - o^2)^2 \\ & - ((a^2 - 1)(n + l + v)^2 + 1 - x^2)^2 \\ & - (16(a^2 - 1)r^2(n + l + v)^4 + 1 - u^2)^2 \\ & - \left(((a + u^2(u^2 - a))^2 - 1)(n + 4d(n + l + v))^2 \right. \\ & \left. + 1 - (x + cu)^2 \right)^2 - ((a^2 - 1)l^2 + 1 - m^2)^2 \\ & - (p + l(a - n - 1) + b(2a(n + 1) - (n + 1)^2 - 1) - m)^2 \\ & - (q + (n + l + v)(a - p - 1) \\ & + s(2a(p + 1) - (p + 1)^2 - 1) - x)^2 \\ & - (z + pl(a - p) + t(2ap - p^2 - 1) - pm)^2 \\ & - (16(k + 1)^3(k + 2)(n + 1)^2 + 1 - f^2)^2 \\ & - (k + 1 + ia - i - l)^2 \\ & \left. - (4g + k + 10 - y(k + 2)(k + 4))^2 \right). \end{aligned}$$

Proof: Let $k \geq 1$. Then $k + 1$ is a younger twin prime if and only if there exists $y \geq 0$ such that $4(k! + 1) + k + 1 = y(k + 1)(k + 3)$. This holds if and only if there exist $g, y \geq 0$ such that $k! = g + 1$ and $4(g + 2) + k + 1 = y(k + 1)(k + 3)$. Now Theorem 2 is used to introduce a diophantine representation of $k! = g + 1$, hence g is replaced by $g + 1$ in all equations. (This allows $g \geq 0$. Theorem 2 states $g \geq 1$ as an assumption.) To avoid double use of the variable y equation (7) of Theorem 2 is cancelled and by virtue of this equation y is replaced by $n + l + v$ in the equations (1)–(6) and (8)–(14). Adding $4(g + 2) + k + 1 = y(k + 1)(k + 3)$ to the set of equations thus obtained results in a diophantine representation of the set of younger twin-primes. Using a simple modification of H. Putnam's construction [16] yields the polynomial as stated in the theorem. The variable k has been replaced by $k + 1$ to allow $k \geq 0$.

Duplication of all equations which contain k gives a similar, though slightly longer equation for all twin-primes. Of course, any diophantine representation of the prime numbers which uses one of the factorials (as [12]) can be employed in the above manner. In his paper [1] P. A. Clement also exhibits characterisations for

prime triples $p, p + 2, p + 6$ and quadruples $p, p + 2, p + 6, p + 8$. They may be employed to yield diophantine representations for these sets, analogous to Corollary 5.

As there are many different diophantine representations of a given recursively enumerable set, the question of the simplest arises. As a measure of complexity the degree, the number of unknowns, the number of additions and multiplications, the polynomial's length or a combination of these may be considered. Employing an idea of T. Skolem, the degree of a diophantine equation can always be lowered to 4 (and thus the degree of any Putnam-polynomial to 5). As for the number of unknowns the situation is much more complicated. It is known, however, that any recursively enumerable set can be represented using 9 unknowns [7], and this leads to 10 variables in the Putnam-polynomial. These numbers are of theoretical interest for sets of prime numbers like Mersenne-primes and Fermat-primes whose cardinality is not known. Any finite set $A = \{a_1, \dots, a_n\}$ may be represented

$$x \in A \Leftrightarrow (x - a_1)(x - a_2) \dots (x - a_n) = 0,$$

thus a single variable suffices. On the other hand, an infinite set of prime numbers cannot be represented with less than two unknowns.

Concerning prime numbers the shortest known polynomial is given in [8], whereas in [13] a Putnam-polynomial in ten variables is constructed. The sets of Mersenne-primes and Fermat-primes can be represented in six unknowns and seven variables in the Putnam-polynomial [6]. Using the methods in [7] and [15] Fermat's last theorem can be represented in 13 unknowns. Recently the Chinese mathematician Sun, Zhi-wei [20], [21] has reduced this number to 10.

ACKNOWLEDGMENT. The author feels indebted to the referee; following his suggestions has led to a considerable improvement in the present paper.

REFERENCES

1. P. A. Clement, Congruences for sets of primes, *Amer. Math. Monthly* 56 (1949), 23–25.
2. M. Davis, Hilbert's tenth problem is unsolvable, *Amer. Math. Monthly* 80 (1973), 233–269.
3. M. Davis, Yu. V. Matijasevič, J. Robinson, Hilbert's tenth problem. Diophantine equations: Positive aspects of a negative solution, in *Mathematical Developments Arising from Hilbert Problems, Proc. Sympos. Pure Math.* 28, Dekalb, 1974, AMS, Providence, 1976, pp. 323–378.
4. M. Davis, H. Putnam, J. Robinson, The decision problem for exponential diophantine equations, *Ann. Math.* 74 (1961), 425–436.
5. R. W. Gosper Jr., Table of the Simple Continued Fraction for π and the Derived Decimal Approximation, *Math. Comp.* 31 (1977), 1044.
6. J. P. Jones, Diophantine representation of Mersenne and Fermat primes, *Acta Arithmetica* 35 (1979), 209–221.
7. ———, Universal diophantine equation, *J. Symb. Logic* 47 (1982), 549–571.
8. J. P. Jones, D. Sato, H. Wada, D. Wiens, Diophantine representation of the set of prime numbers, *Amer. Math. Monthly* 83 (1976), 449–464.
9. S. Lang, H. Trotter, Addendum to “Continued fractions of some algebraic numbers”, *J. reine u. angew. Math.* 267 (1974), 219–220.
10. Yu. I. Manin, *A Course in Mathematical Logic*, Springer-Verlag, New York-Heidelberg-Berlin, 1977.
11. Yu. V. Matijasevič, Enumerable sets are diophantine, *Soviet Math. Doklady* 11 (1970), 354–358.
12. ———, Diophantine representation of the set of prime numbers, *Soviet Math. Doklady* 12 (1971), 249–254.
13. ———, Primes are nonnegative values of a polynomial in 10 variables, *J. Soviet Math.* 15 (1981), 33–44.
14. ———, The Riemann hypothesis from a logician's point of view, in *Number Theory Proc. 1st Conf. Canadian Number Theory Association, Banff*, 1988, de Gruyter, Berlin-New York, 1990, pp. 387–400.

15. Yu. V. Matijasevič, J. Robinson, Reduction of an arbitrary diophantine equation to one in 13 unknowns, *Acta Arithmetica* 27 (1975), 521–553.
16. H. Putnam, An unsolvable problem in number theory, *J. Symb. Logic* 25 (1960), 220–232.
17. P. Ribenboim, *The Book of Prime Number Records*, Springer-Verlag, New York-Berlin-Heidelberg-London-Paris-Tokyo, 1988.
18. J. Robinson, Existential definability in arithmetic, *Trans. Amer. Math. Soc.* 72 (1952), 437–449.
19. C. Smoryński, *Logical Number Theory* 1, Springer-Verlag, Berlin-Heidelberg-New York-London-Paris-Tokyo-Hong Kong-Barcelona-Budapest, 1991.
20. Sun, Zhi-wei, Jones's work on Hilbert's tenth problem, *Advances in Math. (Beijing University)* (to appear).
21. ———, A new relation combining theorem and its application, *Zeitschrift für Math. Logik und Grundlagen der Math.* (to appear).

Department of Mathematics
University of Vienna
Strudlhofgasse 4
A-1090 Vienna
AUSTRIA

THE 1ST CONGRESS

The meeting at Zurich, August 9th–11th, of the International Congress of Mathematicians was in every way a success. More than two hundred members took part. America sent seven representatives, including, however, three Cambridge graduates, now transplanted to Pennsylvania, Professors Harkness, Morley and Charlotte Scott. The greatest mathematician in the world, Sophus Lie, was not expected; and the greatest French mathematician, Poincaré, though down for a speech, did not come; but the actual program was particularly rich and interesting.

It is very noteworthy that the Congress was divided into five sections: (1) Arithmetic and Algebra; (2) Analysis, and Theory of Functions; (3) Geometry; (4) Mechanics and Mathematical Physics; (5) History and Bibliography.

—*American Mathematical Monthly* 4, (1897) p. 229.

Recurrence of Simple Random Walk in the Plane

Terence R. Shore and Douglas B. Tyler

1. INTRODUCTION. To picture a simple random walk in one dimension imagine a person (a random walker) standing at the origin of the real number line. The person repeatedly tosses a coin (which may be biased) and moves successively to the right or left according as the coin shows heads or tails. Does the person ever return to the origin and if so with what probability? The answer is that the person returns to the origin with probability one if the coin is fair, and with probability less than one if the coin is biased.

In this note we consider simple random walk in the plane and show that if the walk is fair then the probability of a return to the origin is also one. This was first proved by Polya in 1921 [1]. (In the same paper Polya showed that simple random walk in \mathbb{R}^3 returns to the origin with probability less than one.) Fair simple random walks in one and two dimensions are said to be recurrent because in each case if the walk returns to the origin once with probability one, it must return infinitely often to the origin with probability one.

The standard proof of Polya's result, as found in texts such as Bailey [2], Feller [3], and Rozanov [4], uses Stirling's approximation of $n!$ at a crucial point in the proof. In section three we use an elliptic integral to express the probability that the random walk returns to the origin. Then in section four we prove Polya's result by using the elliptic integral instead of Stirling's formula. The elliptic integral which we use appeared in Polya's 1921 paper, but was not used there to prove recurrence.

In sections five and six we use the elliptic integral to express the probability, less than one, that certain biased random walks return to the origin. The probability of a return to the origin for these walks can then be computed accurately and easily using Gauss' arithmetic-geometric mean (AGM) method to evaluate the elliptic integral.

2. SIMPLE RANDOM WALK IN THE PLANE—DESCRIPTION, PRELIMINARIES, AND SUMMARY OF RESULTS. For simple random walk in the plane, step k is a random variable X_k that can assume the possible values, right = $(1, 0)$, left = $(-1, 0)$, up = $(0, 1)$, and down = $(0, -1)$ with probabilities p_1 , p_2 , p_3 , and p_4 , respectively. We consider the class of walks in which $p_1 p_2 = p_3 p_4$ because this condition allows us to make use of Spitzer's method [5, p. 89] to show that the components of X_k in the $\vec{i} - \vec{j}$ and $\vec{i} + \vec{j}$ directions are statistically independent, thereby allowing the probability of a return to the origin to be expressed as an elliptic integral. The sequence $\{X_k\}$ is also assumed independent, i.e., for any n

and any finite sequence x_1, x_2, \dots, x_n in the set $M = \{\text{left, right, up, down}\}$,

$$P(X_1 = x_1 \text{ and } X_2 = x_2 \text{ and } \dots \text{ and } X_n = x_n) = \prod_{k=1}^n P(X_k = x_k).$$

Such a sequence $\{X_k\}$ of random variables, is said to be independent and identically distributed (i.i.d. for short). The random walk is the sequence $\{S_n\}$ in which, for each $n \geq 1$, $S_n = \sum_{k=1}^n X_k$ is the position after n steps.

We will be concerned with two sequences of probabilities $\{u_{2n}: n \geq 0\}$ and $\{f_{2n}: n \geq 1\}$ and their generating functions, defined as follows. $u_0 = 1$ and, for each $n \geq 1$, $u_{2n} = P(S_{2n} = (0, 0))$. Thus u_{2n} is the probability that the walk is back at the origin after $2n$ steps. (It cannot be at the origin after an odd number of steps.) For the second sequence, for $n \geq 1$,

$$f_{2n} = P(S_2 \neq (0, 0), S_4 \neq (0, 0), \dots, S_{2n-2} \neq (0, 0), S_{2n} = (0, 0))$$

is the probability that the walk returns to the origin for the first time after $2n$ steps. The generating function of $\{u_{2n}: n \geq 0\}$ is $U(x) = \sum_{n=0}^{\infty} u_{2n} x^{2n}$ and the generating function of $\{f_{2n}: n \geq 1\}$ is $F(x) = \sum_{n=1}^{\infty} f_{2n} x^{2n}$. These series converge at least in the interval $|x| < 1$.

The probability of at least one return to the origin is defined to be

$$F(1) = \sum_{n=1}^{\infty} f_{2n}.$$

$F(x)$ and $U(x)$ are related by the identity (see, e.g., Bailey [2], p. 18)

$$F(x) = 1 - \frac{1}{U(x)},$$

which follows from the fact that for each $n \geq 1$,

$$u_{2n} = u_0 f_{2n} + u_2 f_{2n-2} + \dots + u_{2n-2} f_2$$

so that

$$F(1) = 1 - \frac{1}{\lim_{x \rightarrow 1^-} U(x)}.$$

It is the function $U(x)$ that we express in terms of the elliptic integral

$$I(k) = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2(\theta)}}, \quad 0 < k < 1$$

which is finite for $0 < k < 1$ because the integrand is bounded.

In section 4 we show that if the walk is fair then $F(1) = 1$, i.e. the walk is recurrent. This follows from the fact that $\lim_{x \rightarrow 1^-} I(x) = \infty$.

In section 5 we consider the biased walk in which $p_1 p_2 = p_3 p_4$ and not all of p_1, p_2, p_3 , and p_4 are $1/4$. We write the probability of a return to the origin, less than one, as an elliptic integral, and approximate the probability using Gauss' AGM method in some examples.

Finally, in section 6, we show that if the x - and y -coordinates change independently, and behave as simple random walks on the line, then after a 45-degree rotation of axes and scale change, the walk is simple and satisfies $p_1 p_2 = p_3 p_4$. Thus the probability of a return to the origin can be computed as in section 5.

3. ELLIPTIC INTEGRAL REPRESENTATION OF $U(x)$. In his text [5, p. 89], Spitzer gives us a technique that simplifies the computation of the sequence $\{u_{2n}; n \geq 0\}$. He writes X_k in coordinate form as $X_k = (A_k, B_k)$ and introduces the random vector $Y_k = (A_k - B_k, A_k + B_k)$. Since $S_{2n} = (0, 0)$ is equivalent to $\sum_{k=1}^{2n} A_k = \sum_{k=1}^{2n} B_k = 0$ it follows that $S_{2n} = (0, 0)$ if and only if both $\sum_{k=1}^{2n} (A_k - B_k) = 0$ and $\sum_{k=1}^{2n} (A_k + B_k) = 0$.

One easily checks that for each k : $A_k - B_k$ and $A_k + B_k$ are independent; $P(A_k - B_k = 1) = p_1 + p_4$, $P(A_k - B_k = -1) = p_2 + p_3$, $P(A_k + B_k = 1) = p_1 + p_3$, and $P(A_k + B_k = -1) = p_2 + p_4$. It follows by independence of the sequences $\{A_k - B_k\}$ and $\{A_k + B_k\}$ that

$$\begin{aligned} u_{2n} &= P(S_{2n} = (0, 0)) \\ &= P\left(\sum_{k=1}^{2n} (A_k - B_k) = 0\right) P\left(\sum_{k=1}^{2n} (A_k + B_k) = 0\right) \\ &= \binom{2n}{n} (p_1 + p_4)^n (p_2 + p_3)^n \binom{2n}{n} (p_1 + p_3)^n (p_2 + p_4)^n \\ &= \left(\frac{2n}{n}\right)^2 z^n, \end{aligned}$$

where $z = (p_1 + p_4)(p_2 + p_3)(p_1 + p_3)(p_2 + p_4) \leq 1/16$.

The generating function for the sequence u_0, u_2, u_4, \dots is thus

$$\begin{aligned} U(x) &= \sum_{n=0}^{\infty} \left(\frac{2n}{n}\right)^2 z^n x^{2n}, \quad |x| < 1 \\ &= \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{2^{2n}} (16z)^n x^{2n} \frac{2}{\pi} \int_0^{\pi/2} \sin^{2n}(\theta) d\theta \\ &= \frac{2}{\pi} \int_0^{\pi/2} \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{2^{2n}} (16zx^2 \sin^2(\theta))^n d\theta \\ &= \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - 16zx^2 \sin^2(\theta)}} \\ &= I(4\sqrt{z}x). \end{aligned}$$

In the second and fourth equalities above we have used Wallis' integral and the binomial theorem, namely

$$\frac{\binom{2n}{n}}{2^{2n}} = \frac{2}{\pi} \int_0^{\pi/2} \sin^{2n}(\theta) d\theta \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{2^{2n}} x^n = \frac{1}{\sqrt{1-x}}, \quad |x| < 1.$$

Spitzer's approach showing that $\{A_k - B_k\}$ and $\{A_k + B_k\}$ are independent works if and only if the steps, right, left, up, and down have probabilities that satisfy $p_1 p_2 = p_3 p_4$.

4. FAIR SIMPLE RANDOM WALK RETURNS TO THE ORIGIN WITH PROBABILITY ONE. For a fair walk, $p_i = 1/4$, for $i = 1, 2, 3, 4$, and $z = 1/16$, so that, for $|x| < 1$, we have from section 3 that

$$U(x) = I(x) = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - x^2 \sin^2(\theta)}}.$$

At this point we would like to substitute 1 for x since the integrand would then be $\sec(\theta)$. However the integral is improper at $\pi/2$. Instead we show that $I(x) \rightarrow \infty$ as x approaches 1 from the left.

$$\begin{aligned} I(x) &= \frac{2}{\pi} \int_0^1 \frac{du}{\sqrt{1 - x^2 u^2} \sqrt{1 - u^2}}, \quad \text{using } u = \sin(\theta), \quad du = \cos(\theta) d\theta \\ &\geq \frac{2}{\pi} \int_0^1 \frac{du}{1 - x^2 u^2} \\ &= \frac{1}{\pi x} \log\left(\frac{1+x}{1-x}\right) = \frac{2}{\pi x} \operatorname{arctanh}(x). \end{aligned}$$

Thus $\lim_{x \rightarrow 1^-} U(x) = \infty$ and $F(1) = 1$ which establishes the recurrence of $\{S_n; n \geq 1\}$ for fair simple random walk.

5. FOR BIASED SIMPLE RANDOM WALK THE PROBABILITY OF AT LEAST ONE RETURN IS LESS THAN 1. By biased simple random walk we mean simple random walk in which not all p_i are $1/4$. One way to show that $\{S_n; n \geq 1\}$ is not recurrent in this case is to use the Strong Law of Large Numbers to show that with probability one S_n can return to the origin at most finitely many times, and thus the random walk cannot be recurrent. However, since our goal is to present these results at the undergraduate level (we have done this in two classes) we prefer to use the more elementary approach of generating functions and in addition get an integral representation of the probability of at least one return to the origin.

Using the identity relating $U(x)$ and $F(x)$ and letting x approach 1 from the left we have

$$F(1) = 1 - \frac{1}{I(4\sqrt{z})},$$

where $z = (p_1 + p_4)(p_2 + p_3)(p_1 + p_3)(p_2 + p_4)$. $I(4\sqrt{z})$, which is finite since $k = 4\sqrt{z} < 1$, can be approximated accurately using a numerical integration technique but more quickly using Gauss' AGM procedure. (See Kellogg [6], p. 60 for the latter.) For example, if $p_1 = p_3 = 3/8$ and $p_2 = p_4 = 1/8$, then

$$\begin{aligned} I(4\sqrt{z}) &= I(\sqrt{3}/2) \\ &= \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - \frac{3}{4} \sin^2(\theta)}} \\ &= \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{a_0^2 \cos^2(\theta) + b_0^2 \sin^2(\theta)}}, \end{aligned}$$

where $a_0 = 1$, $b_0 = 1/2$. For each n , replace a_n with

$$a_{n+1} = \frac{a_n + b_n}{2}$$

and b_n with $b_{n+1} = \sqrt{a_n b_n}$. At each iteration the integral remains constant, and since $\{a_n\}$ and $\{b_n\}$ converge to a common limit of approximately .7284, $F(1)$, the probability of at least one return to $(0, 0)$, is approximately $1 - 1/(1/.7284) = .2716$. Since $\{a_n\}$ and $\{b_n\}$ converge rapidly (with about the speed of Newton's method), $F(1)$ can be easily approximated. In this example $a_1 = 3/4$, $b_1 = 1/\sqrt{2}$, $a_2 = .7285$, $b_2 = .7282, \dots$, etc.

For a second example let $p_1 = .01$, $p_2 = .81$, $p_3 = .09$, and $p_4 = .09$. Then $z = .0081$ and $F = 1 - 1/I(.36) = .0338$.

For a final example, monotonicity and continuity of $I(x)$ together with the Intermediate Value Theorem ensure the existence of a unique $p > 1/4$ for which if $p_1 = p_3 = p$ and $p_2 = p_4 = 1/2 - p$, the probability of at least one return is exactly $1/2$. To four places $p \doteq .2939$.

6. INDEPENDENT COORDINATES. For simple random walk in the plane, a step in the horizontal direction precludes a step in the vertical direction and vice versa. What happens if the x - and y -coordinates change independently? Suppose that at each step, the change in x is ± 1 with probabilities α_1 and β_1 with $\alpha_1 + \beta_1 = 1$, and the change in y is ± 1 with probabilities α_2 and β_2 with $\alpha_2 + \beta_2 = 1$. Each step is then one of the four possibilities $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$. Now if we look at the possible positions of this walk, rotate the coordinate system clockwise by 45 degrees, and change the scale by dividing by $\sqrt{2}$ (see Figure 1), we have simple random walk with $p_1 = \alpha_1\beta_2$, $p_2 = \beta_1\alpha_2$, $p_3 = \alpha_1\alpha_2$, and $p_4 = \beta_1\beta_2$. Note that $p_1p_2 = p_3p_4$ and that $z = \alpha_1\beta_1\alpha_2\beta_2 \leq 1/16$. If $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 1/2$ then the walk is recurrent. If not, $4\sqrt{z} < 1$, and $U(1) = I(4\sqrt{z})$, so that $F = 1 - 1/I(4\sqrt{z})$ can be approximated as in section five using Gauss' AGM method.

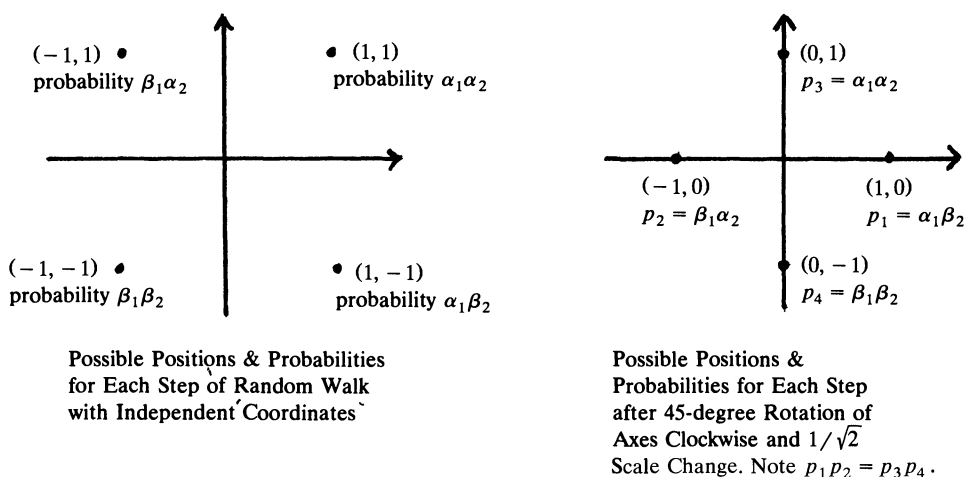


Figure 1

Remark. For random walks in which the step variables X_k are more complex than those we consider here, recurrence means that for an arbitrary neighborhood of the origin, the random walk must return infinitely often to that neighborhood.

ACKNOWLEDGMENTS. The authors thank the referees for their many helpful comments and suggestions. We also thank the referees for suggesting that we investigate random walks with independent coordinates, the subject of section 6.

REFERENCES

1. G. Polya, Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt in Strassennetz, *Mat. Ann.*, 84, 1921.
2. N. Bailey, *The Elements of Stochastic Processes*, Wiley, New York, 1964.
3. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, Wiley, New York, Third Edition, 1968.
4. Y. Rozanov, *Probability Theory: A Concise Course*, Dover, New York, 1969.
5. F. Spitzer, *Principles of Random Walk*, Springer-Verlag, New York, 1964.
6. O. Kellogg, *Foundations of Potential Theory*, Ungar, New York, 1970.

Mathematics Department, NSM A132
California State University
Dominquez Hills
Carson, CA 90747

The MONTHLY draws the attention of mathematicians to the excessive charges for publications of the Cambridge University Press made by the present American agent, The Macmillan Company. A single illustration will suffice.

The third edition of Whittaker and Watson's *A Course of Modern Analysis* was published at 40 shillings (1921, 31); the American agent's price is \$12.50. Hence by ordering from a London bookseller, and paying the duty, a saving of at least \$3.00 on the purchase of this single volume could be effected.—There is no duty for books ordered for college and public libraries. The American Branch of the Oxford University Press appears to count more definitely on the ignorance of purchasers in the United States. To illustrate: H. Hilton's *Plane Algebraic Curves*, 1920, was published at 28 shillings (about \$5.60 at the present rate of exchange); the price of the American Branch is \$12.60!

—*American Mathematical Monthly* 28, (1921) p. 218–219.

Pick's Theorem

Branko Grünbaum and G. C. Shephard

Some years ago, the Northwest Mathematics Conference was held in Eugene, Oregon. To add a bit of local flavor, a forester was included on the program, and those who attended his session were introduced to a variety of nice examples which illustrated the important role that mathematics plays in the forest industry. One of his problems was concerned with the calculation of the area inside a polygonal region drawn to scale from field data obtained for a stand of timber by a timber cruiser. The standard method is to overlay a scale drawing with a transparency on which a square dot pattern is printed. Except for a factor dependent on the relative sizes of the drawing and the square grid, the area inside the polygon is computed by counting all of the dots fully inside the polygon, and then adding half of the number of dots which fall on the bounding edges of the polygon. Although the speaker was not aware that he was essentially using Pick's formula, I was delighted to see that one of my favorite mathematical results was not only beautiful, but even useful. (From DeTemple [1989].)

The discoverer of the theorem in question, Georg Alexander Pick, was born in 1859 in Vienna, and died around 1943 in the Theresienstadt concentration camp. He made significant contributions to analysis and differential geometry. The theorem we are concerned with was first published in 1899 [15]. It became widely known through Steinhaus' delightful book [18].

Pick's theorem concerns lattice polygons ("geoboard polygons"), that is, polygons with all vertices at points of the square unit lattice L , see Figure 1. The original form of the theorem concerns simple polygons, whose edges do not cross one another. (More formally, a polygon is simple if its edges have no mutual intersections other than those of adjacent edges at the common vertices.) The theorem asserts that the area of a simple lattice polygon P is given by the expression

$$i + b/2 - 1,$$

where i is the number of lattice points in the interior of P , and b is the number of lattice points on the boundary of P , that is, points which are either vertices of P or relatively interior points of edges of P . Many proofs of Pick's Theorem are known, see, for example, [1], [2], [3], [6], [7], [10], [11], [12], [14]; there are various generalizations: to more general polygons [9], [15], [19], to lattices other than the square lattice [4], [5], and to higher-dimensional polyhedra [13], [16], [17], [20].

In this paper we shall extend Pick's theorem to more general lattice polygons, by allowing multiple intersections, and even overlapping, of the edges. We shall

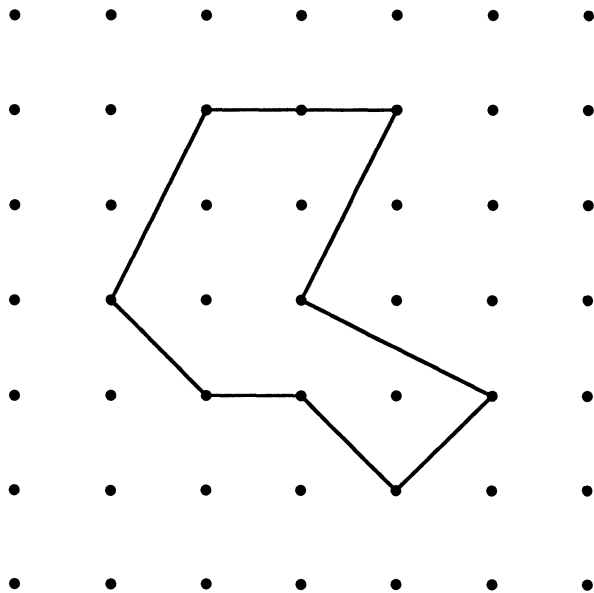


Figure 1. A simple lattice 8-gon P illustrating the classical form of Pick's Theorem. Here $i = 4$ is the number of lattice points in the interior of P , $b = 9$ is the number of lattice points on the boundary of P , and

$$A(P) = i + \frac{b}{2} - 1 = \frac{15}{2}$$

is the area of P .

make use of results on rotation numbers, winding numbers and tangent numbers of such polygons P ; a brief account of the necessary definitions and facts concerning these numbers will be given here, but for more details, examples, and proofs of some of our assertions, the reader should consult [8].

1. BASIC DEFINITIONS. By an *abstract polygon* or *n-gon* Q we mean an ordered sequence (V_1, \dots, V_n) of n distinct symbols V_1, \dots, V_n , called the *vertices* of Q . Adjacent pairs $(V_1, V_2), (V_2, V_3), \dots, (V_n, V_1)$ (where the subscripts are taken mod n) are called the *edges* of Q , and two sequences which differ only by a cyclic permutation of the symbols are regarded as identical. Thus Q has a definite *orientation* and the edge (V_i, V_{i+1}) is said to be *oriented* or *directed* from V_i to V_{i+1} .

A *lattice polygon* is any embedding of an abstract polygon and its edges in the plane, such that the following conditions hold:

(i) The image of each V_i is a point of the square unit lattice L . Without confusion we may continue to denote the image of V_i by the same symbol, and to call it a vertex of the polygon.

(ii) Each edge $(V_i, V_{i+1}) \pmod{n}$ is represented by a straight line segment connecting the image points V_i, V_{i+1} in the plane. In the diagrams it is convenient to denote the direction of the edge by an arrow.

An example of a lattice polygon is shown in Figure 2(a). If we impose two additional restrictions, namely

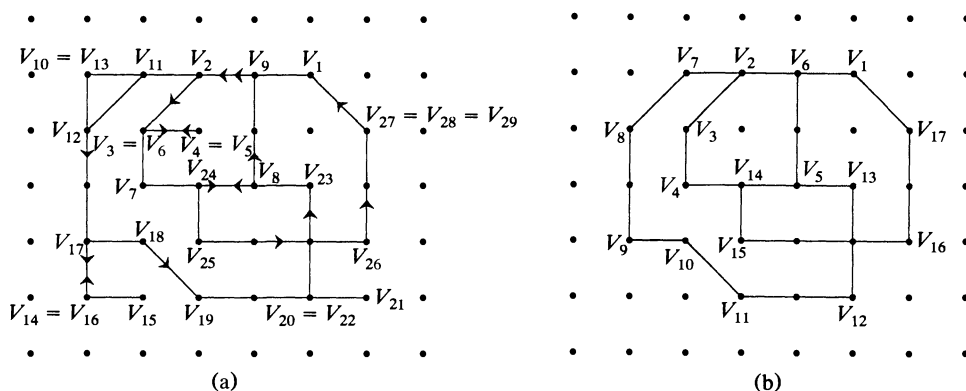


Figure 2. (a) An example of a lattice polygon (29-gon) with multiple vertices ($V_3 = V_6$, $V_4 = V_5$, $V_{10} = V_{13}$, $V_{14} = V_{16}$, $V_{20} = V_{22}$, $V_{27} = V_{28} = V_{29}$) and whiskers ($V_9, V_{10}, V_{11}; V_{12}, V_{13}, V_{14}; V_{14}, V_{15}, V_{16}; V_3, V_4, V_6$ also becomes a whisker after the multiple vertex $V_4 = V_5$ has been removed, and V_{12}, V_{14}, V_{17} becomes a whisker after the removal of whiskers V_{12}, V_{13}, V_{14} and V_{14}, V_{15}, V_{16}). (b) The polygon that results from shaving that shown in (a); it is a 17-gon.

(iii) no two consecutive vertices $V_i, V_{i+1} \pmod n$ map onto the same lattice point, and

(iv) two edges with a common vertex do not overlap (that is, the polygon has no “whisker”),

then we say that the polygon is *shaven*. Examples of shaven polygons appear in Figures 2(b) and 3. (The numbers attached to the lattice points in the latter figure will be explained in the next section.) Figure 3 shows some of the possibilities for multiple intersections and overlaps of edges that are not excluded by conditions (i) to (iv).

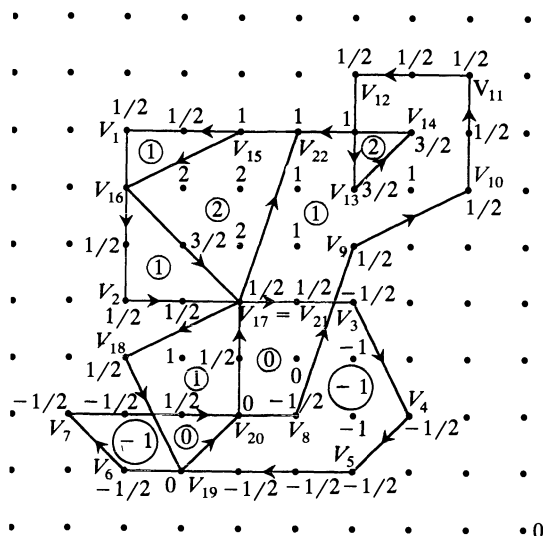


Figure 3. A general lattice polygon P with 22 vertices. The winding numbers of the various cells are shown by circled numbers. The rotation number of P is 2, and the indices of the lattice points are indicated. (The index of every point in the exterior is 0, and is not marked.) The area of P is 18 and the sum of all the indices is $i(P) = 20$, in agreement with the assertion of Theorem 1 that $A(P) = i(P) - r(P)$.

We shall usually denote a lattice polygon (shaven or otherwise) by the letter P . The complement $\mathbb{E}^2 \setminus P$ of P in the plane consists of a finite number of connected open regions called the *cells* of P . All are bounded except for one, which is called the *exterior* of P .

Now let X be a given point of the plane, which does not belong to P , and let $R(X)$ be a ray (a closed half-line) with endpoint X which does not pass through any vertex of P . Then for each edge E_j we define

$$\omega(R(X), E_j) = \begin{cases} 0 & \text{if } R(X) \text{ does not intersect } E_j, \\ 1 & \text{if } E_j \text{ crosses } R(X) \text{ in a counterclockwise} \\ & \text{direction as viewed from } X, \\ -1 & \text{if } E_j \text{ crosses } R(X) \text{ in a clockwise direction.} \end{cases}$$

The *winding number* $w(P, R(X))$ of P with respect to $R(X)$ is defined as $\sum_j \omega(R(X), E_j)$ summed over all the edges of P . It can be shown that $w(P, R(X))$ depends *only* on the endpoint X of $R(X)$, and *not* on the particular ray $R(X)$ that was used. In fact, this even applies to rays that pass through vertices of P if the definition is suitably modified. In view of this we may use the notation $w(P, X)$ unambiguously. Further, it can also be shown that if X and Y belong to the same cell of P , then $w(P, X) = w(P, Y)$. Hence we can define the winding number $w(P, C)$ of a cell C with respect to P as the winding number of any point in the cell. In Figure 3 the winding numbers of the cells are indicated. These are the same winding numbers which are well known from calculus for their rôle in defining the area enclosed by curves with selfintersections.

The *area* $A(P)$ of P is defined as $\sum_j w(P, C_j) |A(C_j)|$ summed over all the cells C_j of P . Here $A(C_j)$ is the (usual) elementary area of the polygonal region C_j . Figure 3 serves to illustrate the calculation of the area of the polygon P . The area of a polygon can be positive, negative or zero. Since the winding number $w(P, C_j)$ of a cell changes its sign if we reverse the orientation of P (that is, reverse the order of the vertices in the definition of the polygon), the same is true for the area of P .

Next, we need the concept of “rotation number” (sometimes called “tangent winding number”) of a polygon. Throughout we shall use the *absolute system* of angle measure, in which a complete counterclockwise turn of 2π radians has value 1. At a vertex V_j of a shaven polygon P let W lie on the extension of (V_{j-1}, V_j) beyond V_j . Then the signed angle $\angle WV_jV_{j+1}$ (which necessarily satisfies $-\frac{1}{2} < \angle WV_jV_{j+1} < \frac{1}{2}$) is called the *deflection* $d(V_j)$ of P at V_j . It is easy to show that $r(P) = \sum_j d(V_j)$, with summation over all the vertices of P , is necessarily an integer, called the *rotation number* of P . For the polygon P of Figure 3 we have $r(P) = 2$, as indicated in the caption. It should be observed that the rotation number is only defined for shaven polygons, since the definition of deflection is not applicable at multiple vertices or whiskers.

To facilitate the formulation of the next definition, we note that if a vertex V_j of the lattice polygon lies on a ray $R(X)$ with endpoint X , then the two edges (V_{j-1}, V_j) and (V_j, V_{j+1}) which meet at V_j can lie in six different configurations with respect to the ray, see Figure 4. If the edges lie on different sides of $R(X)$ (cases (a) and (b)) we say that $R(X)$ *cuts* the polygon P at V_j . In the four other cases we say that P is *tangent* to $R(X)$ at V_j . In (c) and (f) we say that the tangency is *concordant* since the directions induced on $R(X)$ by the edges (V_{j-1}, V_j) , (V_j, V_{j+1}) are consistent with that on $R(X)$ oriented away from X . In (d) and (e) we say that the tangency is not concordant.

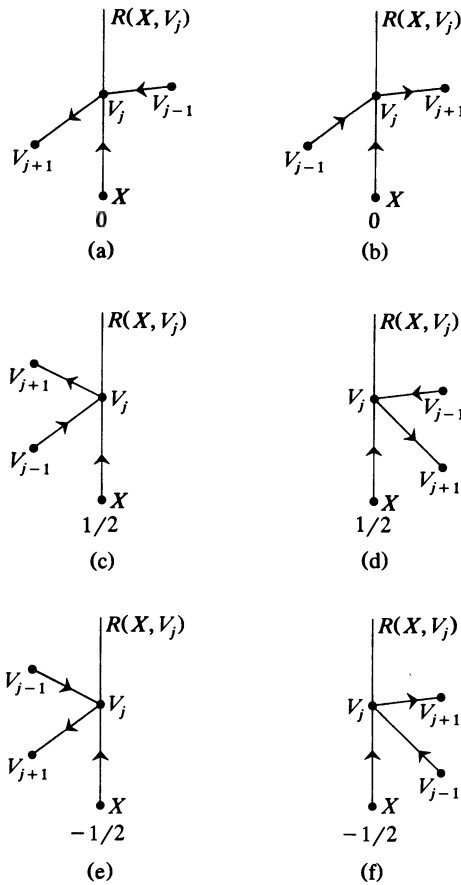


Figure 4. The possible positions of edges (V_{j-1}, V_j) and (V_j, V_{j+1}) relative to ray $R(X)$, and the contributions to the tangent number $t(X, P)$.

Let now X be any point of the plane which does not lie on a line containing an edge of a lattice polygon P , and consider a closed ray $R(X)$ with endpoint X that passes through a vertex V_j of P . Then we define

$$\tau(R(X), V_j) = \begin{cases} 0 & \text{if } R(X) \text{ cuts } P \text{ at } V_j \text{ (cases (a) and (b) of Figure 4),} \\ \frac{1}{2} & \text{if the edges } (V_{j-1}, V_j) \text{ and } (V_j, V_{j+1}) \text{ lie to the left} \\ & \text{of } R(X) \text{ and the tangency is concordant (case (c)),} \\ & \text{or they lie to the right of } R(X) \text{ and the tangency is} \\ & \text{not concordant (case (d)),} \\ -\frac{1}{2} & \text{in all other cases ((e) and (f) in Figure 4).} \end{cases}$$

For fixed X let $t(P, X) = \sum_j \tau(R(X), V_j)$, where the sum is over all the vertices of P . It can be shown that $t(P, X)$ is necessarily an integer; it is known as the *tangent number* of X with respect to P . (This definition differs slightly from the one given in [8], but is equivalent to it.) We require the following important property of $t(P, X)$ (see [8]):

If X lies in the exterior of P , then $t(P, X) = r(P)$.

$R(X)$. (In the case of multiple intersections, the contributions due to each arc of P are added.) From the definition, it is clear that if $X \notin P$ then $i(P, X) = w(P, x)$.

Lemma. *The value of $i(P, X)$ depends only on the lattice point X and the polygon P and not on the ray $R(X)$ chosen to define it.*

Proof: Consider changes in the value of $i(P, X)$ that occur as $R(X)$ is rotated in a counterclockwise direction about X . Let Z be any point on $R(X)$ such that the open line segment $]X, Z[$ lies entirely in some cell of P . Suppose the initial position of the ray is $R_1(X)$ (see Figure 6) and Z is at Z_1 . It is clear that during the rotation, so long as Z does not cross any edge of P , then the fact that $w(P, Z)$ remains constant shows that $i(P, X)$ does so also.

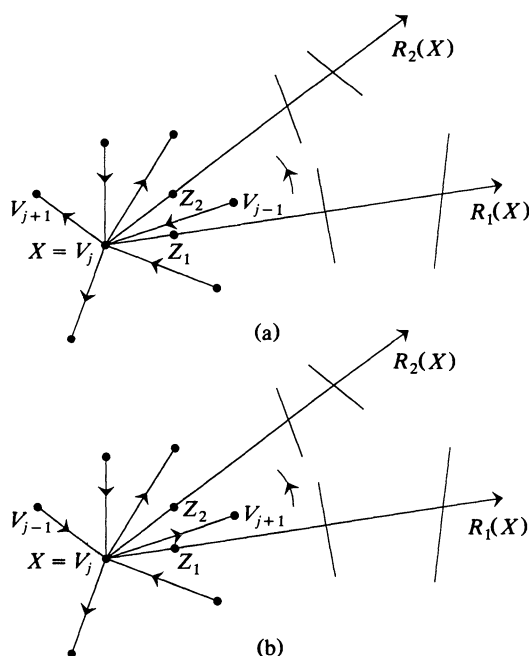


Figure 6. Diagrams illustrating the proof of Lemma. The intersection of P with $R(X)$ at points other than X are shown schematically at right in each of the diagram.

Now let $R(X)$ cross an edge directed towards X (see Figure 6(a)) to position $R_2(X)$ with Z moving to Z_2 . The change in $i(X, P)$ will be twofold:

(i) the contribution to $i(P, X)$ from the intersections of P with $R(X)$ at points other than X (which equals the winding number of Z with respect to P) will decrease by 1, and

(ii) the contribution to $i(X, P)$ from the intersections of P with $R(X)$ at X will increase by 1. (In the diagram the edges (V_{j-1}, V_j) and (V_j, V_{j+1}) form a non-concordant tangency to $R_1(X)$ at X and so contribute $-\frac{1}{2}$. However, they cut $R_2(X)$ from right to left and so in this new situation contribute $\frac{1}{2}$. The total change in the contribution is $+1$. It is easy to check that in all other cases the same holds.)

On the other hand, if $R(X)$ crosses an edge directed away from X (see Figure 6(b)) then the corresponding changes are $+1$ and -1 respectively. If pairs of edges coincide, then their contributions are added.

Thus it will be seen that in all cases the rotation of $R(X)$ about X does not alter the value of $i(P, X)$ as we have defined it, and so the lemma is proved.

At first sight the definition of $i(P, X)$ may appear somewhat artificial. In fact, as will be seen from Figure 1, in the case of a simple polygon P oriented in a positive (counterclockwise) direction, the index of each lattice point in the interior of P is 1 and the index of each lattice point on the boundary of P is $\frac{1}{2}$. Theorem 1, stated at the beginning of the next section, reduces immediately to the classical form of Pick's Theorem in this case. The complications in the definition of $i(X, P)$ arise because of the need to deal with lattice points that occur at multiple intersections and overlapping edges which may occur in the general lattice polygons which we are considering here.

Finally, we write $i(P) = \sum i(P, X)$, where summation is over all the lattice points X of L . We note that this sum is finite since $i(P, X) = 0$ for all X in the exterior of P .

3. SHAVEN POLYGONS. We begin with the basic theorem from which the more general result (Theorem 2 of the next section) can be derived.

Theorem 1. *Let P be any shaven lattice polygon. Then*

$$A(P) = i(P) - r(P),$$

where $A(P)$ is the area of P , $r(P)$ is the rotation number, and $i(P)$ is the sum of the indices with respect to P of all the lattice points.

Proof: Let O be any point in the plane not belonging to P . We calculate the area of P in the classical way, as follows. Let T_j be the triangle obtained by joining the edge E_j of P to O (that is, T_j is the convex hull of E_j and O with the orientation induced by that of E_j). Then $A(P) = \sum_j A(T_j)$, summation being over all the edges of P , and areas being counted with appropriate signs ($A(T)$ is positive if T is oriented in a counterclockwise direction, and negative if T is oriented clockwise.)

In the present context we take O as a lattice point in the exterior of P , and apply the classical form of Pick's Theorem to find the area of each triangle. All that is needed for the proof is an investigation as to how the indices of the lattices points and rotation numbers change when two triangles T_1 and T_2 corresponding to adjacent edges (V_{j-1}, V_j) and (V_j, V_{j+1}) are welded together along their common boundary $[O, V_j]$, see Figure 7. Consider, to begin with, the case where T_1 and T_2 are oriented positively (case (a)). The index of a lattice point in the relative interior of T_1 is 1 and of a point on its boundary is $\frac{1}{2}$. Also the rotation number is 1, and so $A(T_1) = i(T_1) - 1 = i(T_1) - 2i(T_1, O)$. Similarly $A(T_2) = i(T_2) - 2i(T_2, O)$. After welding the triangles together we obtain a quadrilateral Q with vertices V_{j-1}, V_j, V_{j+1}, O and we note that the index of each lattice point with respect to Q is the sum of the indices assigned to the point by consideration of T_1 and T_2 *except for the points V_j and O* . Since $i(T_1, V_j) = i(T_2, V_j) = i(Q, V_j) = \frac{1}{2}$, we must *subtract* $\frac{1}{2}$ from the indices of V_j and O , and then

$$A(Q) = i(Q) - 1 = i(Q) - 2i(Q, O).$$

The other cases (b) to (f) in Figure 7 are dealt with similarly. In case (c), for example, $i(T_1, V_j) = -\frac{1}{2}$, $i(T_2, V_j) = i(Q, V_j) = \frac{1}{2}$ and $i(T_1, O) = -\frac{1}{2}$, $i(T_2, O) = i(Q, O) = \frac{1}{2}$. Hence, after summing the indices we must *add* $\frac{1}{2}$ to those of V_j and

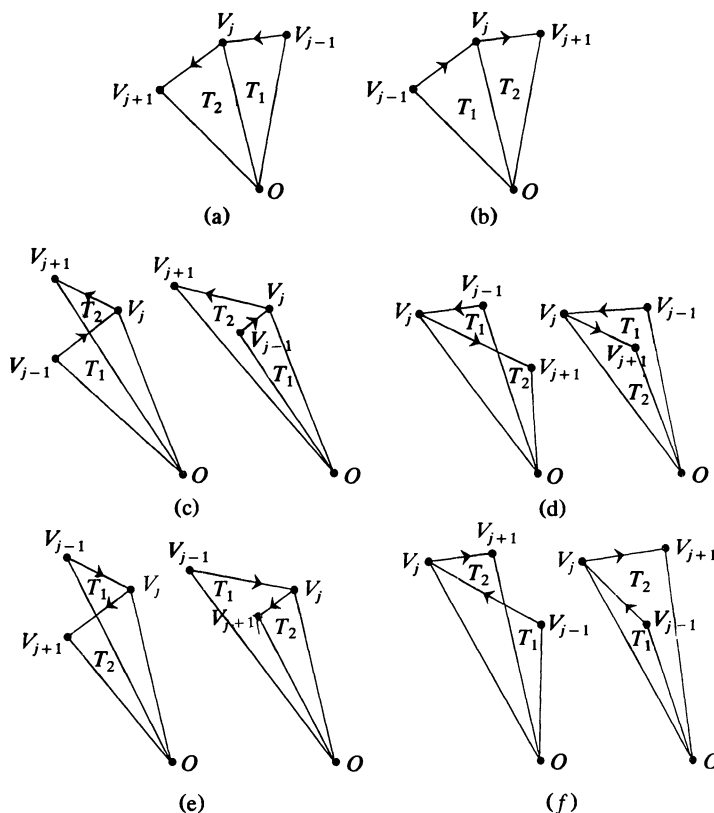


Figure 7. The possible configurations of oriented triangles $OV_{j-1}V_j$, OV_jV_{j+1} with an edge OV_j in common. These are used in the proof of Theorem 1.

O . However, the relation

$$A(Q) = i(Q) - 1 = i(Q) - 2i(Q, O)$$

continues to hold. It does so also in the other cases: for (b) (c) and (d) we must *add* $\frac{1}{2}$ to the indices of V_j and O after amalgamation of the triangles, and we must *subtract* $\frac{1}{2}$ in cases (a), (e) and (f).

We build up P triangle by triangle, making the necessary modifications to the indices at the ends of the common edges. By natural extension of the above, at each stage X except the final one,

$$A(X) = i(X) - 2i(X, O).$$

When the triangle corresponding to the last edge of P is adjoined, and the modifications to the indices are applied as described above, we obtain the indices $i(P, X)$ of all the lattice points X with respect to P , *except that at O there will be an index i^* which is not equal to $i(P, O)$* . Thus the sum of all the indices will be $i(P) + i^*$, and the area will be given by

$$A(P) = (i(P) + i^*) - 2i^*.$$

To complete the proof we need only evaluate i^* . Let t^+ and t^- be the numbers of positively and negatively oriented triangles which meet at O , and $t_a, t_b, t_c, t_d, t_e, t_f$ be the numbers of vertices V_j of P at which the adjacent edges lie

in the configurations (a), (b), (c), (d), (e), (f) of Figure 7, respectively. Then the above construction shows that

$$\begin{aligned} i^* &= \frac{1}{2}(t^+ - t^- - t_a + t_b + t_c + t_d - t_e - t_f) \\ &= \frac{1}{2}((t^+ - t_a) - (t^- - t_b) + t_c + t_d - t_e - t_f). \end{aligned}$$

Now $t^+ - t_a$ is the number of connected chains of edges of P which are oriented counterclockwise viewed from O , and $t^- - t_b$ is the number of connected chains of edges of P which are oriented clockwise viewed from O . As O is exterior to P these must be equal, and the first two terms cancel leaving

$$i^* = \frac{1}{2}(t_c + t_d - t_e - t_f).$$

Comparing Figures 7 and 4, we see that $i^* = i(P, O) = r(P)$ by the result quoted above. Thus

$$A(P) = i(P) - r(P),$$

and the theorem is proved.

Now let $\mathbb{P} = \{P_1, P_2, \dots, P_n\}$ be a family of shaven polygons, that is to say, a finite set of such polygons. We define

$$\begin{aligned} r(\mathbb{P}) &= \sum_{j=1}^n r(P_j), & A(\mathbb{P}) &= \sum_{j=1}^n A(P_j), \\ i(\mathbb{P}, X) &= \sum_{j=1}^n i(P_j, X), & i(\mathbb{P}) &= \sum_{j=1}^n i(P_j). \end{aligned}$$

Then, by additivity, Theorem 1 immediately implies the following:

Corollary. *If \mathbb{P} is any family of shaven lattice polygons then*

$$A(\mathbb{P}) = i(\mathbb{P}) - r(\mathbb{P}).$$

4. GENERAL LATTICE POLYGONS. We began, in Section 1, by defining a general lattice polygon, but our main result (Theorem 1) applies only to shaven polygons. It is of some interest to consider whether this restriction is necessary. For *any* lattice polygon P the indices of the lattice points are uniquely defined as in Section 2, hence $i(P)$ is well defined. The area $A(P)$ is also uniquely determined, as already explained. However, the rotation number $r(P)$ is indeterminate if P has whiskers, or pairs of adjacent vertices which coincide; therefore to obtain a result analogous to Theorem 1 which applies to general polygons we need to obtain some number which plays a rôle analogous to $r(P)$.

To do this, we obtain a shaven polygon P' from the general polygon P by

(a) removing each whisker. More precisely, if V_{j-1}, V_j, V_{j+1} is a whisker we delete the vertex V_j and replace the edges (V_{j-1}, V_j) and (V_j, V_{j+1}) by either the edge (V_{j-1}, V_{j+1}) or by the two coinciding adjacent vertices $V_{j-1} = V_{j+1}$.

(b) removing coincident adjacent vertices. Thus if $V_j = V_{j+1} = \dots = V_{j+s}$ we remove V_{j+1}, \dots, V_{j+s} and replace the edge (V_{j+s}, V_{j+s+1}) by (V_j, V_{j+s+1}) .

Either operation can produce further whiskers, and operation (a) can produce multiple vertices. However, it can be shown easily that repeated applications of the operations lead eventually to a shaven polygon P' . Even though the polygon P' is in some cases not uniquely determined by P , the rotation number $r(P')$ is independent of the particular P' obtained, and we may therefore define $r(P)$ to have the value $r(P')$. With this convention we obtain the following form of Pick's theorem.

Theorem 2. Let P be any lattice polygon, $A(P)$ its area, $i(P)$ the sum of the indices at the lattice points, and $r(P)$ the rotation number as defined above. Then

$$A(P) = i(P) - r(P).$$

Proof: This is immediate from the observation that $A(P) = A(P')$, $i(P) = i(P')$, $r(P) = r(P')$, and from Theorem 1 applied to the shaven polygon P' .

5. FINAL REMARKS. A major difference between our treatment of Pick's Theorem and that of earlier papers is that *oriented* polygons are used here. If suitable orientations are introduced our main theorem implies all known variants of Pick's Theorem as it relates to plane lattice polygons. It does not, of course, include the three-dimensional version of [13], [16], [17], nor the results of [4], [5] concerning the "hexagonal lattice" (which is not a lattice in the usual terminology) or Archimedean tilings.

Since the index of a lattice point, the rotation number of a polygon, and the area of a polygon are invariant under affine transformations of determinant 1, it follows that our theorems are invariant under such transformations. Hence it includes results relating to polygons whose vertices lie at lattice points of the equilateral triangular lattice, see [5]. For "polygons of higher genus" such as those of [14, Figures 5, 6, 7], after suitable orientations of the edges are introduced, the results follow from our Theorem 1 and its corollary.

Many generalizations of Pick's Theorem introduce, as one of the variables in the formula for the area, the Euler characteristic of the polygon (or the polygonal region). Since the area depends on the orientation of the polygon, whereas the Euler characteristic does not, it is not an appropriate variable to use for polygons or families of polygons that may have regions with winding numbers other than 0 or 1. If the only winding numbers are 1 or 0, then it is trivial to orient the polygons and deduce the results in the literature from ours.

The same remark applies to the setting in [9]. Here a further simplification is possible. Each of the "zweiseitige Randstrecken" ("two-sided boundary edges") corresponds to two overlapping edges with opposite orientations. Even if these are not whiskers they may be removed in a manner exactly analogous to that described in (a) above. This reduces the problem of determining the area, in the examples given, to an application of the corollary to Theorem 1.

It was pointed out to us by Prof. Rolf Schneider that our results could be reformulated in the following way: For each lattice point X in the plane assign an index $j(X, P)$ that equals the sum of the angular lengths of arcs of a small circle centered at X that are contained in P , with the appropriate signs and multiplicities. Then $A(P) = \sum j(X, P)$, where the summation is over all lattice points in the plane.

ACKNOWLEDGMENT. The authors are indebted to Prof. Duane DeTemple for helpful comments, and for copies of various relevant articles, including Pick's original paper.

REFERENCES

1. H. S. M. Coxeter, *Introduction to Geometry*, Wiley, New York 1969.
2. D. DeTemple, Pick's formula: A retrospective, *Mathematics Notes from Washington State University*, Vol. 32, Nos. 3–4 (November 1989).
3. D. DeTemple and J. M. Robertson, The equivalence of Euler's and Pick's Theorems, *Math. Teacher* 67 (1974), 222–226.

4. R. Ding, K. Kolodziejczyk and J. R. Reay, A new Pick-type theorem on the hexagonal lattice, *Discrete Math.* 68 (1988), 171–177.
5. R. Ding and J. R. Reay, The boundary characteristic and Pick's theorem in the Archimedean planar tilings, *J. Combinat. Theory* A44 (1987), 110–119.
6. W. W. Funkenbusch, From Euler's formula to Pick's formula using an edge theorem, *Amer. Math. Monthly* 81 (1974), 647–648.
7. R. W. Gaskell, M. S. Klamkin and P. Watson, Triangulations and Pick's theorem, *Math. Mag.* 49 (1976), 35–37; comments by J. Staib and R. A. Gibbs, *ibid.* pp. 104–105, 158.
8. B. Grünbaum and G. C. Shephard, Rotation and winding numbers for polygons and curves, *Trans. Amer. Math. Soc.* 322 (1990), 169–187.
9. H. Hadwiger and J. M. Wills, Neuere Studien über Gitterpolygone, *J. reine angew. Math.* 280 (1975), 61–69.
10. G. Haig, A 'natural' approach to Pick's theorem, *Math. Gazette* 64 (1980), 173–177.
11. R. Honsberger, *Ingenuity in Mathematics*, New Mathematical Library, volume 23, Math. Association of America, Washington, D.C. 1970, pp. 27–31.
12. A. C. F. Liu, Lattice points and Pick's theorem, *Math. Mag.* 52 (1979), 232–235.
13. I. G. MacDonald, The volume of a lattice polyhedron, *Proc. Cambridge Philos. Soc.* 59 (1963), 719–726.
14. I. Niven and H. S. Zuckerman, Lattice points and polygonal area, *Amer. Math. Monthly* 74 (1967), pp. 1195–1200. Reprinted in *Selected Papers in Geometry*, A. K. Stehney et al., eds., Math. Assoc. of America, 1979, pp. 149–153.
15. G. Pick, Geometrisches zur Zahlenlehre, *Sitzungber. Lotos* (Prague) 19 (1899), 311–319.
16. J. E. Reeve, On the volume of lattice polyhedra, *Proc. London Math. Soc.* (3) 7 (1957), 378–395.
17. J. E. Reeve, A further note on the volume of lattice polyhedra, *J. London Math. Soc.* 34 (1959), 57–62.
18. H. Steinhaus, *Mathematical Snapshots*. Oxford Univ. Press, New York 1969.
19. D. E. Varberg, Pick's theorem revisited, *Amer. Math. Monthly* 92 (1985), 584–587.
20. J. M. Wills, Kugellagerungen und Konvexgeometrie, *Jahresber. Deutsch. Math.-Verein*, 92 (1990), 21–46.

Department of Mathematics
University of Washington
Seattle, WA 98195

University of East Anglia
Norwich, NR4 7TJ
ENGLAND

John Wiley & Sons recently published a volume entitled "Theory and applications of finite groups," consisting of three parts. Part I, written by Professor G. A. Miller, consists of 192 pages and is entitled "Substitution and abstract groups"; Part II, written by Professor H. F. Blichfeldt, consists of 86 pages and is entitled "Finite groups of linear homogeneous transformations"; Part III, written by Professor L. E. Dickson, consists of 103 pages and is entitled "Applications of finite groups." The work is dedicated to Camille Jordan, and is the first treatise on group theory written by American mathematicians.

—*American Mathematical Monthly* 23, (1916) p. 317.

Mathematics for Liberal Arts Students

Albert W. Briggs, Jr.

I have been developing a course for liberal arts students for many years and it is time to report on these efforts. I found that the process of developing this course changed the way I teach in all courses but in ways that make teaching harder, not easier. And I have had some insights into the reasons why studying mathematics usually doesn't teach students to think. First some history.

Near the end of the spring 1975 semester, a student who had taken finite math from me the semester before hailed me as I walked across the campus and asked, "Hey Doc, guess what?" "What?" I responded. He said, "It's been a semester since I took the course and I haven't seen a Markov chain yet." The first thing that popped into my mind was, "Yeah, kid, and I bet you never do." But I never shared that with him. Instead, I responded with some statements about how generally useful Markov chains were, and about how he would know what was being talked about if anyone who worked for him used them. I didn't give him any of the "training the mind" justifications because I knew him and I knew he would never buy that. He grudgingly accepted what I said and walked away, but I knew I had to change the course.

I remembered the saying that your education was what was left after you threw away your notes and forgot everything you ever knew. In his case, what would be left? I couldn't find much. He and many other students had taken the course with utilitarian goals in mind. But except for some linear programming and the elementary algebra it required, I doubt they ever used anything in the course while they were undergraduates. And once they left? It seemed to me that the small number of people who might have contact with a small number of these topics at some distant time in no way justified all this effort by so many people.

Groping for a solution to this dilemma, I used *Mathematics, A Human Endeavor*, by Harold R. Jacobs, in both semesters of the '75-'76 year. I began reading Polya's *How to Solve It*, *Mathematics and Plausible Reasoning*, and *Mathematical Discovery*, as well as Wayne Wickelgren's *How to Solve Problems*. I noticed that mathematicians often referred to each other as "investigators," so in the spring semester I assigned a term paper, which was to be an "investigation" of one of the questions on a list I handed out. This meant that if the student could not answer the question, then the investigation would be the partial results obtained. If the question could be answered, then any other questions that arose as a result also had to be pursued. This "Written Investigation" is an assignment I have made ever since.

One student wrote a paper on, "Is it possible to place 9 chips on a table in such a way that they form 10 lines with exactly 3 chips on each line?" Interpreting "chips" and "table," to be "points" and "plane," respectively, he looked at a 3 by

3 square array, found the obvious 8 lines and at first thought the answer might be “No.” But then he noticed that there was no necessity to arrange the points in a square, and moving them around found an arrangement that gave 9 lines, exactly 3 points per line. He couldn’t find an arrangement that gave 10 lines, so he decided again that the answer was “No.” He tried to find a proof and thought he had succeeded. However, one step in it bothered him, so he constructed an example to test that step and there in front of him was an arrangement that gave 10 lines. Now that he knew the answer was “Yes,” he set out to prove that no more than 10 lines could be formed, but could not. (The last time I checked, neither could anyone else.) Still, he was convinced that 10 was the maximum. Then he asked, “How many lines can be put through 16 points, exactly 4 to a line?” He found what he thought was the maximum, then extended this to 25 points 5 to a line, and so on up to 64 points 8 to a line. Formulating this as finding the maximum number of lines possible through N^2 points, exactly N to a line, he found a formula that would give all the maxima he had found. He tried, but could not prove the conjecture. (Some time later, I believe I saw the same conjecture published in the *Journal of Recreational Mathematics*, but I can’t locate it now.) It occurred to me later that what he had done was just what a mathematician who didn’t know any more than he did would have done. After that, it also occurred to me that long after he had forgotten what his conjecture was, or that he had worked on the problem, or that he even had taken the course, he could still use his mind like a mathematician. That was just what I wanted. Eventually, the thought occurred to me that I should try to get these benefits by going directly at them instead of studying some topic for the sake of utility or some other reason and hoping these benefits would occur as a side effect.

Despite this one good outcome, the semester was a disaster. I thought that if I just conducted the course without stating the aims, much as I did in other courses, the students would catch on. But they didn’t. They kept expecting a course something like the ones they had known and they expected a gut, but this was neither. I got some of the most scathing student evaluations I have ever seen. I found that with a course as different as this, I had to state the real goals of the course early and refer to them often, neither of which I had done that semester.

In an attempt to eliminate these difficulties, in the fall of 1976 I gave the students an expanded course description. It also provided the faculty with some unintended humor. It has been lost, but I remember that most of the sentences in it began something like, “This is not a course in . . .,” or “This course is not intended to . . .”. That I had said what the course wasn’t, but not what it was, had to be pointed out to me. It illustrated very well the huge struggle I was having with myself. Namely, the other courses I taught were similar enough to those I had taken so that by proceeding instinctively I achieved the explicit goals of teaching the topics in question and probably achieved any implicit goals as well without my thinking about them or even knowing precisely what they were. But in this course I didn’t believe in a topic to be learned or taught. Hence a period of denial, specifically, denying topics, was a psychological stage through which I had to go. Once I went through it, I became disenchanted with Jacobs and all other such books, because they focused on the results to be obtained, classifying them by topic, not on the mental processes by which they were obtained. So from the fall of 1977 through the fall of 1984, I used no text. Instead, I began using material from Polya, Wickelgren, and many recreational mathematics sources, in an effort to teach the heuristics of problem solving. And I began to use material from *Proofs*

and Refutations by Lakatos in an effort to show how theory building could occur. The course was becoming my answer to the question, “What math course should students take who want to broaden their educations but don’t need to know any specific topic, such as statistics or calculus?”

There are two points that need to be understood about this stage of the process. First, I finally realized that one reason I was having so much trouble achieving the goals of the course was that I had never thought them through in sufficient detail to be able to design a course to achieve them. Second, I found that it is one thing to write a draft of the aims of the course, but it is quite another to achieve those aims while conducting the class in a mode as intensely interactive as was required.

Eventually, I came to the conclusion that the specific goals for the course are that the students should learn:

1. to raise questions concerning and make conjectures about objects and ideas of mathematical interest,
2. to test those conjectures,
3. to settle definitively those questions and conjectures or at least make progress toward such a settlement,
4. to raise more questions as a result of this process,
5. to read with enough precision to understand the questions and investigations of others,
6. to write with enough precision to communicate their investigations to others.

I refer to them often in class to supply the context for particular tasks I set and events that occur. For example, we would start with a question and try to answer it. If the discussion of it reveals that what is known or what is desired has not been established, I would not only raise those questions but point out that I am raising them, that learning to raise them is part of the course and that they are useful for other problems too. As another example, if a student guesses that some statement might be true, I might point out that a guess has been made, that we must decide whether to test it with examples or try to prove it and that both are what we are learning how to do. (Incidentally, I have found that students take to forming guesses and testing them with examples like ducks take to water. But getting them to prove their guesses, or to deduce things from their guesses in an effort to explore, is killingly hard. It is so hard that I have to wonder whether deduction, as we practice it in mathematics, is as natural a human tendency as we suppose.) Sometimes the discussion reveals that we may be headed up a blind alley. Then, following ideas in Alan Schoenfeld’s *Mathematical Problem Solving*, I will act as a manager and ask what we will do with the results when we get them, as well as pointing out that I’m doing so and why. Similarly, if it appears that some belief, e.g., that every math problem can be done in 5 minutes or less, is interfering with our investigation, I will point that out to illustrate the fact that beliefs can hurt, or help, investigations. In consciously considering such a belief, we may be able to change it. With regard to writing, I emphasize writing up questions tackled in class. I begin this by showing how I would write up something we did in class, then require them to write up problems which we have discussed until everyone claims to have understood them. The audience for this writing is to be a student who is “good at math,” but who doesn’t know anything about this particular question. Since we concentrate on the process so hard, many students want to give me a blow-by-blow description of how the problem was done instead of telling me what

the question is, what the answer is and how we know it is the answer. I try to sell this as exercise in yet another pattern, a pattern of communication in a modern style. Some students immediately see the beauty in presenting only the results, but others find it flavorless.

A question I ask myself in an effort to guide my classroom behavior is, "If a perceptive mathematician were to drop in on this class after the goals had been discussed and without knowing anything about the course, could that person tell that we were trying to achieve something like the aims above?" If the answer is, "No, you might be thought to be doing a course in recreational mathematics," then I'm probably doing something wrong. Specifically, I'm probably concentrating too hard on the math and not enough on the processes by which it was done, or I'm treating the math as an end in itself, not as an example of how to do math. Of course, I can't monitor myself this way at the same time I spontaneously interact with the class. It must be done sequentially. But the monitoring eventually produced changes in my spontaneous interactions.

That the sex of the mathematician in the question above cannot be determined brings out a minor, but interesting, observation. I found that if I state all the questions in a deliberately sex-neutral way, ("Person A," "Team B", etc.,) then when students are discussing the problem with me in my office, male students tend to use masculine pronouns and female students tend to use feminine pronouns. I have no proof that their assumptions help them, but I have kept that style ever since, despite its awkwardness. By now it is almost a habit.

In the process of developing this course, I found that I had new insights into other courses and changed the way I teach all of them. I now include the topic-specific material we normally teach in the larger contexts of solving problems, or of discovering mathematics. In fact, I can't refrain from doing so. As a simple example, consider the precalculus problem, "Show that the quadrilateral whose vertices are $(-4, 1)$, $(0, -2)$, $(6, 6)$, and $(2, 9)$ is a parallelogram." In discussing it, I might ask, "Do we understand what is being asked?" This might be followed by, "What is a parallelogram?" This in turn might be followed by, "What could we know about a quadrilateral that would convince us it is a parallelogram?" If I insist that the students respond to this line of questioning, they are likely to request that I, "stop asking all these questions and tell us how to do the problem." That is because the context within which they have studied math does not include self-consciously working backward nor self-consciously exploring the meanings of the terms. The steps to solve a problem either occur to them unself-consciously, or they don't. If I treat these questions as rhetorical, answering them myself, the students don't complain because they will find out how to do the problem if they wait. But they don't learn to raise questions themselves. The context within which they have learned mathematics gives them no way to attack new problems even in the same mathematical topic, let alone in other mathematical topics, nor in quantitative areas still further afield. The only way around this I have found is to discuss the context explicitly just as I have learned, painfully, to do with liberal arts students. This takes time. I also have observed that, like the liberal arts students, mathematics students who have not learned spontaneously to operate in a more general context, have great difficulty learning how to do so. It is as if these students and I are in different Piagetian stages, a stage I admit I entered in the process of developing the liberal arts course and my classroom behavior in it. (I don't recognize this as a stage Piaget described exactly and I suspect he understated the number of stages by at least an order of magnitude.) As with all such

barriers, this difference has made it harder to teach. It is also the reason I think math courses don't usually teach students to think. The only gains most experience are an increased ability to acquire new mathematical resources and to re-acquire resources that have been forgotten.

Department of Mathematics and Computer Science
Washington College
Chestertown, MD 21620

Studies in the Theory of Numbers. By Leonard E. Dickson. The University of Chicago Science Series, 1930. x + 230 pages. \$4.00.

This important volume has two claims to distinction: it contains an amazing number of new results in the theory of quadratic forms; and it represents a systematic treatment of the arithmetic theory of quadratic forms, starting from first principles. Either one of these accomplishments by itself would entitle the author and his students and collaborators (Arnold Ross, Gordon Pall, A. Oppenheim) to the lasting gratitude of all interested in the theory of numbers; the combination makes the book of quite outstanding value.

It would seem to call for some explanation why a systematic treatment *ab ovo* of an apparently well developed field such as the arithmetic theory of quadratic forms should be hailed as a noteworthy achievement. It will probably be to many readers, as it was to the reviewer, a shock to learn that in spite of the eminence of the mathematicians who have contributed to the theory (Gauss, Seeber, Smith, Zolotareff, Markoff, Frobenius, Minkowski, Eisenstein; to mention only some of those no longer living) and in spite of the very large number of textbooks on theory of numbers, we have no satisfactory exhaustive treatment of this field. In particular, the volumes of Bachmann "Die Arithmetik der Quadratischen Formen" are shown to be in important respects unreliable.

One can but admire the courage of an author who will undertake to rebuild the whole structure rather than to patch up the unsound portions. One can only guess at the amount of labor covered by the modest words of the preface; "It was no small task to write a satisfactory exposition." On the other hand, we know, in this country as well as in Europe, how much the theory of numbers owes to the insistence of Dickson on precision in the statement of theorems and to his uncanny ability to detect, and to mend, unsound arguments; it seems therefore only fair that to him and his students should belong the credit of writing the first reliable treatment of the arithmetic theory.

—*American Mathematical Monthly* 40, (1933) p. 40.

NOTES

Edited by: John Duncan

Sequences with Large Numbers of Prime Values

Ulrich Abel and Hartmut Siebert

It is not known whether there are polynomials of degree greater than 1 with integer coefficients representing infinitely many primes for integer argument. However, for any N , Sierpinski [4] proved that c can be chosen such that $x^2 + c$ represents at least N primes. Garrison [2] generalized this to $x^n + c$.

In the theorem below we show that the result is true for every sequence (a_n) of positive integers whose counting function

$$A(x) = \sum_{\substack{n \\ a_n \leq x}} 1$$

satisfies certain conditions. As a corollary we infer that Sierpinski's result holds for arbitrary polynomials of degree greater than 1 with integer coefficients.

Our argument of proof depends on counting the number of solutions of certain inequalities and shows that no arithmetical properties of polynomials are needed other than their rate of growth.

This makes an idea that was already implicit in Sierpinski's argument, completely transparent.

Theorem. *Let (a_n) be a sequence of pairwise different positive integers with the counting function $A(x)$.*

If

$$\limsup_{x \rightarrow +\infty} \frac{A(x)}{\log x} = +\infty, \quad (1)$$

then for every integer N there exists a positive integer $c = c(N)$ such that $a_n + c$ represents more than N primes.

If

$$\limsup_{x \rightarrow +\infty} \frac{A(2x) - A(x)}{\log x} = +\infty, \quad (2)$$

then for every integer N there exists a positive integer $c = c(N)$ such that $a_n - c$ represents more than N primes.

Proof: By Sylvester's version of the Chebyshev inequalities ([6], [7], see also [1], (1.7), p. 555) there exists an absolute positive constant R such that simultaneously,

$$0.9 \leq \pi(x) \frac{\log x}{x} \leq 1.1 \quad (x \geq R) \quad (3)$$

and

$$\frac{\log x}{\log(2x)} > \frac{8}{9} \quad (x \geq R). \quad (4)$$

Then for every natural number N there exists, by (1), an integer $x > R$ with

$$A(x)/\log x > 4N. \quad (5)$$

Let $Z(x)$ be the number of solutions of the inequality

$$0 < p - a_n \leq 2x,$$

which we can estimate by $Z(x) \geq [\pi(2x) - \pi(x)]A(x)$. Now, by (3), (4) and (5),

$$\begin{aligned} [\pi(2x) - \pi(x)]A(x) &\geq \left(0.9 \frac{2x}{\log 2x} - 1.1 \frac{x}{\log x}\right)A(x) \\ &= \left(1.8 \frac{\log x}{\log 2x} - 1.1\right) \frac{A(x)}{\log x} x > 0.5 \cdot 4 \cdot N \cdot x = 2xN. \end{aligned}$$

Therefore we can find an integer $c \in [1, 2x]$ with $c = p - a_n$ for more than N primes p and indices n . This yields the first part of Theorem 1.

If now $Z(x)$ denotes the number of solutions of the inequality

$$0 < a_n - p \leq 2x,$$

we get the estimation $Z(x) \geq [A(2x) - A(x)]\pi(x)$. Similarly one can show that $Z(x) > 2xN$ for some integer x , and the assertion follows.

Corollary. Let $P(x) = \sum_{k=0}^m b_k x^k$ be a polynomial of degree $m \geq 1$ with integer coefficients b_k ($k = 0, 1, \dots, m$) and $b_m > 0$. Then, for every positive N there exists an integer $c = c(N)$, such that $P(x) + c$ is a prime for more than N integers x . The sign of $b_0 + c$ can be prescribed to be positive or negative.

The corollary contains Sierpinski's and Garrison's results as special cases.

Proof of the corollary. Because of $b_m > 0$ there exists a positive integer n_0 such that P is increasing for $x \geq n_0$ and

$$\beta x^m \leq P(x + n_0) - b_0 \leq 1.5\beta x^m \quad (x \geq 0)$$

with a certain constant $\beta > 0$.

We show that the sequence (a_n) with

$$a_n = P(n + n_0) - b_0$$

satisfies (2) of the theorem. Note that (2) implies (1).

$$\begin{aligned} A(2x) - A(x) &= \sum_{\substack{n \\ a_n \leq 2x}} 1 - \sum_{\substack{n \\ a_n \leq x}} 1 \\ &\geq \sum_{\substack{n \\ 1.5\beta n^m \leq 2x}} 1 - \sum_{\substack{n \\ \beta n^m \leq x}} 1 \\ &= \left[\left(\frac{4x}{3\beta} \right)^{1/m} \right] - \left[\left(\frac{x}{\beta} \right)^{1/m} \right] \\ &\geq Kx^{1/m} \quad (x > x_0) \text{ with a positive constant } K. \end{aligned}$$

Therefore, (2) holds.

We do not know whether our theorem is best possible in the sense that its assertion becomes false when the sequence (a_n) grows such that (1) or (2) are violated. See also a problem recently proposed by S. Golomb [3] in this Monthly.

ACKNOWLEDGMENT. The authors would like to thank the referee who pointed out that in the proof of the theorem it is sufficient to use Chebyshev's inequalities instead of the prime number theorem.

REFERENCES

1. H. G. Diamond, Elementary methods in the study of the distribution of prime numbers, *Bull. AMS* 7 (1982), 553–589.
2. B. Garrison, Polynomials with large numbers of prime values, *Amer. Math. Monthly* 97 (1990), 316–317.
3. S. Golomb, Problem no. 10208, *Amer. Math. Monthly* 99 (1992), 266.
4. W. Sierpinski, Les binomes $x^2 + n$ et les nombres premiers, *Bull. Soc. Royale Sciences Liège*, 33 (1964), 259–260.
5. W. Sierpinski, Elementary theory of numbers, *Polska Akademia Nauk*, Warschau 1964.
6. J. J. Sylvester, On Tchebycheff's theorem of the totality of prime numbers comprised within given limits, *Amer. J. Math.* 4 (1881), 230–247.
7. J. J. Sylvester, On arithmetical series, *Messenger of Math.* (2) 21 (1892), 1–19 and 87–120.

Fachhochschule Giessen-Friedberg
Fachbereich MND
Wilhelm-Leuschner-Strasse 13
D-6360 Friedberg
GERMANY

Regular Simplices in Spaces of Constant Curvature

Horst Martini

1. INTRODUCTION. A well-known theorem states the following: If all 2-faces of a tetrahedron T in Euclidean 3-space R^3 have equal areas, then these 2-faces are congruent triangles (see e.g. Court [6] and Couderc-Ballicioni [5]). With $\{A_1, A_2, A_3, A_4\}$ being the vertex set of T , this means $|A_1A_2| = |A_3A_4|$, $|A_1A_3| = |A_2A_4|$, $|A_1A_4| = |A_2A_3|$. (It should be noted that such tetrahedra are called “isosceles” and that they can be used for characterizing Euclidean motions, cf. Lenz [11]. An interesting characterization of isosceles tetrahedra in R^3 by means of certain centroids was given by Bottema [2].) Among other results, Horváth [9] proved the analogous implication for tetrahedra in 3-dimensional spherical and hyperbolic space; for the hyperbolic case we refer to Bui Van Dung [3; 4] and Lenz-Selényi-Zeitler [12], too.

It seems to have been overlooked that the result of Horváth implies an elementary characterization of regular n -simplices in n -dimensional spherical and hyperbolic space (denoted by S^n and H^n , respectively) for $n \geq 4$: Such a simplex is regular if and only if all its 2-faces have equal areas.

This characterization extends a corresponding result of Frankl-Maehara [8] for Euclidean n -simplices ($n \geq 4$), which therefore is reproved (as a special case of our statements) in an elementary way. For further recent results on simplices in spaces of constant curvature we mention a paper of Dekster-Wilker [7].

2. ISOSCELES TETRAHEDRA IN 3-SPACES OF CONSTANT CURVATURE.

We shall start with an outline of the results from [9]. Horváth proved that the following properties of a tetrahedron T in Euclidean, spherical or hyperbolic 3-space are equivalent:

- (I) The 2-faces of T have equal areas.
- (II) The 2-faces of T are congruent.
- (III) The measures of the stereoangles at the vertices of T are equal.

Here the measure of the stereoangle at a vertex A_1 of the tetrahedron T with vertices A_1, \dots, A_4 is defined as follows: If ω_{ik} denotes the measure of the angle of the faces of T at the edge $A_i A_k$, then the measure of the stereoangle at A_1 is given by

$$\delta_1 := \omega_{12} + \omega_{13} + \omega_{14} - \pi.$$

(It should be noticed that the Euclidean version of (I) \Leftrightarrow (III) goes back to Kármány [10].)

For proving (I) \Rightarrow (II), Horváth used metrical properties of:

- Saccheri quadrangles suitably constructed to the faces of T ,
- the centrally symmetric quadrangles Q_1 , Q_2 , and Q_3 , whose vertices are the midpoints of four edges of T , in each case without one of the three skew pairs of edges,
- the orthogonal projections of the vertex set of T onto the planes spanned by Q_1, Q_2, Q_3 .

Finally, these considerations yield $|A_1 A_2| = |A_3 A_4|$, $|A_1 A_3| = |A_2 A_4|$, $|A_1 A_4| = |A_2 A_3|$ under the assumption (I). The converse implication is trivial. The equivalence of (II) and (III) was derived by congruence arguments with respect to the four trihedra (= unions of 3 faces) at the vertices of T .

3. THE REGULARITY OF n -SIMPLICES ($n \geq 4$). By definition, an n -simplex S ($n \geq 2$) in n -space of constant curvature is regular if its symmetry group acts transitively on all its r -faces for each r between 0 and n (see e.g. Böhm-Hertel [1], §6). It is easy to deduce that S is regular if and only if all its edges have the same length (cf. once more [1]). Using this and the results from [9], we shall prove the announced

Theorem. *For $n \geq 4$, the following properties of an n -simplex S in Euclidean, spherical or hyperbolic n -space are equivalent:*

- (1) The simplex S is regular.
- (2) The 2-faces of S have equal areas.
- (3) The 2-faces of S are congruent.
- (4) The measures of the four stereoangles of each 3-face of S are equal to each other.

Proof: We shall show that the equality of all 2-face areas of an n -simplex in R^n , S^n and H^n ($n \geq 4$) implies the equal length of all its edges. For seeing this, it is sufficient to verify this implication for $n = 4$, since a non-regular n -simplex ($n > 4$) must have some non-regular 4-face. We shall study the edge-skeleton of such a 4-simplex S having a tetrahedron T as one of its 3-faces. The results of Horváth imply that a non-regular tetrahedron T with equal 2-face areas in R^3 , S^3 or H^3 can only have an edge-skeleton of one of the following two types:

$$\begin{aligned} |A_1A_2| = |A_3A_4| &=: \bar{a}, & |A_1A_3| = |A_2A_4| &=: \bar{b}, \\ |A_1A_4| = |A_2A_3| &=: \bar{c}, \end{aligned}$$

or

$$|A_1A_2| = |A_3A_4| =: \bar{a}, \quad |A_1A_3| = |A_2A_4| = |A_1A_4| = |A_2A_3| =: \bar{b}$$

with $\bar{a} \neq \bar{b} \neq \bar{c} \neq \bar{a}$.

Thus, if two congruent edges of T have another length (say \bar{a} from above) than any of the other four edges, they present a pair of skew edges and contain the whole vertex set of T . Now we consider a second 3-face of S , having the vertices A_1 , A_2 , A_4 , and A_5 , say. Then the congruence of all 2-faces of S implies $|A_1A_2| = |A_4A_5| = \bar{a}$. But this contradicts $|A_4A_5| = \bar{a}$ in view of a third 3-face with vertices A_2 , A_3 , A_4 , A_5 . Namely, for this third 3-face $|A_3A_4| = \bar{a}$ implies that in addition only its edge A_2A_5 can have the length \bar{a} .

Hence, assuming equal 2-face areas for S , a 3-face cannot have a pair of edges with another length than the remaining four edges of it, and therefore, S can only have edges of equal length. The converse implication (1) \Rightarrow (2) is trivial, and the equivalences of Horváth show that also (3) and (4) are equivalent to (1).

REFERENCES

1. J. Böhm, E. Hertel, *Polyedergeometrie in n -dimensionalen Räumen konstanter Krümmung*, Deutscher Verlag der Wissenschaften, Berlin 1980.
2. O. Bottema; The centroids of a simplex (Dutch), *Euclides*, Groningen 47 (1971/1972), 206–210.
3. Bui Van Dung, Some properties of equilateral tetrahedra with ideal vertices in the hyperbolic space (Hungarian, Russian Summary), *Mat. Lapok* 32 (1984), 127–135.
4. Bui Van Dung, Some properties of tetrahedra in hyperbolic space (Hungarian, Russian summary), *Mat. Lapok* 32 (1985), 219–228.
5. P. Couderc, A. Ballicioni, *Pemier livre du tétraèdre*, Gauthier-Villars, Paris 1953.
6. N. A. Court, *Modern Pure Solid Geometry*, Macmillan, New York 1935.
7. B. V. Dekster and J. B. Wilker, Simplexes in spaces of constant curvature, *Geometriae Dedicata* 38 (1991), 1–12.
8. P. Frankl and H. Maehara, Simplices with given 2-face areas, *European J. Combin.* 11 (1990), 241–247.
9. J. Horváth, A property of tetrahedra with equal faces in spaces of constant curvature (Hungarian, German summary), *Mat. Lapok* 20 (1969), 257–263.
10. F. Kárteszi, A geometric extremum problem (Hungarian), *Mat. Lapok* 18 (1967), 67–74.
11. H. Lenz, Über einen Satz von J. Lester zur Charakterisierung euklidischer Bewegungen, *Journal of Geometry* 28 (1987), 197–201.
12. H. Lenz, G. Selényi, H. Zeitler: Einige Eigenschaften gleichflächiger Tetraeder in der hyperbolischen Geometrie, Working papers, Pécs-Osijek 3 (1989), 81–89.

*Grossmannstr. 13
D / O-8512 Grossröhrsdorf
GERMANY*

A Simple Example of Little Big Set

John K. Williams

Loosely speaking, the Hausdorff dimension of a set is the right place to measure that set. For s less than the Hausdorff dimension, the Hausdorff s -measure of the set is infinite and for s larger than the Hausdorff dimension the Hausdorff s -measure is zero. This note describes a simple example of a set with Hausdorff s -measure 0, little, and Hausdorff dimension s , big. Specifically we will construct a set with Hausdorff dimension 1 and Hausdorff 1-measure 0 and then show how it can be generalized to give such a set for any s .

To get a feel for Hausdorff dimension let's look at an example, the Cantor set. The Cantor set is defined as follows. Let $E_0 = [0, 1]$ and then define E_j as E_{j-1} with the open middle one third of each interval removed, i.e. $E_1 = [0, 1/3] \cup [2/3, 1]$, $E_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$. Each E_j consists of 2^j intervals of length 3^{-j} . Cantor's set is the set $E = \bigcap_{j=0}^{\infty} E_j$.

The Hausdorff 1-measure is just the same as the 1-dimensional Lebesgue measure, the length of the set. From the construction, one can see that the length of the E_j is $2^j \times (1/3)^j$ and the length of $E = \lim_{j \rightarrow \infty} (2/3)^j = 0$. Now if we change how we measure the length of an interval, $[a, b]$, from $|a - b|$ to $|a - b|^s$, then the length of each E_j becomes $2^j \times (1/3)^{js}$. The length of $E = \lim_{j \rightarrow \infty} (2/3^s)^j$. If $s > (\ln 2 / \ln 3)$, then the limit is 0; if $s < (\ln 2 / \ln 3)$, the limit is infinity; and if $s = (\ln 2 / \ln 3)$, the limit is 1. As we will see below, the Hausdorff dimension of this set is precisely $(\ln 2 / \ln 3)$. Now a definition of Hausdorff dimension [1, p. 7].

Define the diameter of a non-empty subset U of \mathfrak{R}^n as $|U| = \sup\{|x - y| : x, y \in U\}$. If $E \subseteq \bigcup_i U_i$ and $0 < |U_i| \leq \delta$ for each i , we say that $\{U_i\}$ is a δ -cover of E .

For E a subset of \mathfrak{R}^n , s a non-negative number, and $\delta > 0$ define:

$$\mathfrak{H}_{\delta}^s(E) = \inf \sum_{i=1}^{\infty} |U_i|^s,$$

where the infimum is over all (countable) δ -covers $\{U_i\}$ of E . Then the *Hausdorff s -dimensional outer measure* of E is defined as:

$$\mathfrak{H}^s(E) = \lim_{\delta \rightarrow 0} \mathfrak{H}_{\delta}^s(E).$$

The limit exists since \mathfrak{H}_{δ}^s increases as δ decreases but may be infinite. The restriction of \mathfrak{H}^s to the σ -field of \mathfrak{H}^s -measurable sets is called the *Hausdorff s -dimensional measure*.

For each E , $\mathfrak{H}^s(E)$ is non-increasing as s increases from 0 to ∞ (as soon as δ is less than 1, $|U_i|^s$ decreases as s goes from 0 to ∞). Also if $s < t$, then

$$\mathfrak{H}_{\delta}^s(E) \geq \delta^{s-t} \mathfrak{H}_{\delta}^t(E),$$

which implies that if $\mathfrak{H}^t(E)$ is positive, then $\mathfrak{H}^s(E)$ is infinite. There is then a unique value, $\dim E$, called the *Hausdorff dimension* of E , such that

$$\mathfrak{H}^s(E) = \infty \text{ if } 0 \leq s < \dim E \quad \text{and} \quad \mathfrak{H}^s(E) = 0 \text{ if } \dim E < s < \infty.$$

Three observations are immediate. First, if E is a subset of F then the Hausdorff dimension of E is less than or equal to the Hausdorff dimension of F . Secondly, the Hausdorff dimension of \mathbb{R}^n is n . Putting the first two together, the Hausdorff dimension of a subset of \mathbb{R}^n is less than or equal to n . The question addressed in this note is can $\mathfrak{H}^s(E)$ be zero when s is $\dim E$?

If E is \mathbb{R}^1 , then $\dim E$ is 1 and $\mathfrak{H}^1(E) = \infty$. If E is a line segment of length l , then $\dim E$ is again 1 and $\mathfrak{H}^1(E)$ is l . Is there a set E where $\dim E$ is 1 while $\mathfrak{H}^1(E) = 0$? At first it does not seem possible. The idea of Hausdorff dimension is to find a place where the set should be measured, for s smaller than the dimension, the measure is infinite and for s larger, the measure is zero. A slight modification of the Cantor set will lead us to a set which has the desired properties.

One can view the Cantor set as the invariant set for a pair of linear contractions applied to the unit interval. If we define:

$$f_1(x) = (1/3)x \quad \text{and} \quad f_2(x) = (1/3)x + 2/3,$$

Then each E_j above may be defined inductively as $E_0 = [0, 1]$ and $E_j = f_1(E_{j-1}) \cup f_2(E_{j-1})$.

The advantage of viewing the Cantor set in this manner is that we can apply a theorem of Moran [3, Thm. II] which in this context can be stated as follows:

Theorem. *Let $\{f_1, \dots, f_n\}$ be a set of linear contractions, each of which contracts by a factor of w_n . Let E_0 be a set where $f_j(E_0)$ and $f_k(E_0)$ are disjoint for $j \neq k$, $E_j = \bigcup_{i=1}^n f_i(E_{j-1})$, and $E = \bigcap_{j=0}^{\infty} E_j$. Then the Hausdorff dimension of E is s where s is defined by the equation:*

$$\sum_{i=1}^n w_i^s = 1$$

and the Hausdorff s -measure is finite and positive.

Applying this to the Cantor set, we see that the Hausdorff dimension is $\log 2 / \log 3$ and has positive Hausdorff $(\log 2 / \log 3)$ -measure.

We can modify the construction of the Cantor set slightly by taking out the middle $1/m$ th section at each step. This means modifying the linear contractions to be:

$$f_1(x) = \left(\frac{m-1}{2m}\right)x \quad \text{and} \quad f_2(x) = \left(\frac{m-1}{2m}\right)x + \frac{m+1}{2m}.$$

Applying Moran theorem we have the following theorem:

Theorem. *The Hausdorff dimension of \mathcal{C}_m is*

$$s = \log 2 / \log \left(\frac{2m}{m-1} \right)$$

and $\mathfrak{H}^s(\mathcal{C}_m)$ is finite and positive.

Now the set we are looking for is just the union of the \mathcal{C}_m . Set $\mathcal{C} = \bigcup_{j=3}^{\infty} \mathcal{C}_m$.

Theorem. *The Hausdorff dimension of \mathcal{C} is 1 and $\mathfrak{H}^1(\mathcal{C}) = 0$.*

Since the dimension of each of the \mathcal{C}_m is less than one, $\mathfrak{H}^1(\mathcal{C}_m) = 0$. Therefore

$$\mathfrak{H}^1(\mathcal{C}) = \mathfrak{H}^1\left(\bigcup_{m=3}^{\infty} \mathcal{C}_m\right) \leq \sum_{m=3}^{\infty} \mathfrak{H}^1(\mathcal{C}_m) = 0.$$

But the dimension of \mathcal{C} must be greater than or equal to the dimension of each \mathcal{C}_m . Since the dimension of \mathcal{C}_m tends to 1 as m tends to ∞ , the dimension of \mathcal{C} must be one. \square

This construction can be generalized in two directions. First we can construct a Cantor like set of any dimension s between 0 and 1 by pulling out the middle α at each stage where s and α are related by:

$$\alpha = 1 - 2^{1-1/s}.$$

To find a set with dimension s and Hausdorff s -measure 0, find a sequence which converges to s , $\{s_i\}$; construct a Cantor like set for each s_i ; and form the union of these sets.

Secondly we can reach dimensions higher than 1 by pulling squares out the unit square {a Sierpinski Gasket} or pulling cubes out the unit cube in dimension three and above.

Finally, there are other sets with this property. Most notably, a Besicovitch set is a set with zero planar measure, lines in all directions and Hausdorff dimension 2. The exact description and proof of its properties is a little more difficult [1, ch. 7].

ACKNOWLEDGMENT. My thanks to the referee who pointed out that Federer [2] has a similar construction, in greater generality of course.

REFERENCES

1. K. J. Falconer, *The Geometry of Fractal Sets*, New York: Cambridge University Press, 1985.
2. H. Federer, *Geometric Measure Theory*, New York, Springer, 1969.
3. P. A. P. Moran, Additive functions of intervals and Hausdorff Measure' *Proceedings of the Cambridge Philosophical Society*, 42 (1946) 15–23.

*Department of Mathematics, Physics and Computer Science
University of Hartford
West Hartford, CT 06117*

E 402 [1940, 48]. *Proposed by Irving Kaplansky, Harvard University.*

If n , r , and a are positive integers, the congruence $n^2 \equiv n \pmod{10^a}$ obviously implies $n^r \equiv n \pmod{10^a}$. (When such a number n has only a digits, it is called an automorphic number.) For what values of r does $n^r \equiv n \pmod{10^a}$ imply $n^2 \equiv n \pmod{10^a}$?

—*American Mathematical Monthly* 47, (1940) p. 572.

LETTERS

Alternating Matrices

I would like to bring to your attention, concerning the article “On the product of two alternating matrices” (The MONTHLY, 98 (1991), 935–936) by D. Ž. Đoković, the fact that a complete characterization of the products of two alternating (skew-symmetric) matrices has already appeared in the paper “Pairs of alternating forms and products of two skew-symmetric matrices” (Linear Algebra Appl., 63 (1984), 119–132) by R. Gow and T. J. Laffey: The n by n matrix A is such a product if and only if

(1) the elementary divisors of A corresponding to nonzero eigenvalues have even multiplicity, and

(2) the elementary divisors of A corresponding to zero eigenvalue are of the form

$$x^{k_1}, x^{k'_1}, \dots, x^{k_s}, x^{k'_s} \quad \text{or} \quad x^{k_1}, x^{k'_1}, \dots, x^{k_s}, x^{k'_s}, x,$$

where $k'_i = k_i$ or k_{i+1} , $1 \leq i \leq s$.

The main results in Đoković's article are of course easy corollaries of this.

Pei Yuan Wu
Department of Applied Mathematics
National Chiao Tung University
1001 Ta Hsueh Road, Hsinchu
Taiwan, REPUBLIC OF CHINA

Bessel Functions

Readers of the article *Bessel Functions and Kepler's Equation* by Peter Colwell in the January, 1992, *Monthly* may be interested to know that there is a discussion of Kepler's equation in Section 89 of the book *Théorie des Résidues* by H. Laurent (Gauthier-Villars, Paris, 1865). Laurent considers *une équation que l'on rencontre en Astronomie*: $z = x + t \sin z$. In Section 87, he investigates solutions of Lagrange's equation $w = a + t\phi(w)$, described in Colwell's article. There is no discussion, however, of the work of Bessel and Carlini. Laurent's book appears to be one of the first giving an exposition of Cauchy's theory of residues.

Roderick Gow
Department of Mathematics
University College Dublin
Belfield, Dublin 4
IRELAND

Fundamental Theorem

Professor Joseph Bennis's admirably concise and elementary proof of the Fundamental Theorem of Algebra ["Another Proof of the Fundamental Theorem of Algebra", *American Mathematical Monthly*, 99, (May 1992), p. 426] may become still clearer with one ending comment. Some proper subsets of the plane have both a non-empty interior and a boundary consisting of only finitely many points, for example, the plane punctured at the origin. Because the Riemann sphere is compact, however, so must be its image under a continuous map, for instance, a polynomial. As a compact subset of the Riemann sphere, the image is also closed, which excludes the possibility of any isolated point on the boundary of its interior, which such points would puncture. Consequently, from Bennis's proof that the boundary consists of at most finitely many points follows that the boundary is empty. Therefore, the interior of the image is open, non-empty, and has an empty boundary, which means that it covers the entire connected Riemann sphere.

Yves Nievergelt
Department of Mathematics, MS-32
Eastern Washington University
Cheney, WA 99004-2415
ynieverg%ewu@uunet.uu.net

Corrigendum

I write to call attention to a serious typographical error in the paper, "Reliability, Recursion, and Risk," by L. B. Page and J. E. Perry (the *Monthly*, v. 98 #10, Dec. 1991, pp. 937–946).

Because the paper presented a novel but accessible technique for computing probabilities of moderately-complex events, I chose to discuss it in an undergraduate course for computer-science students here at Brock University. Our discussion was greatly complicated by the above-mentioned error; the paper's Figure 5 presents its major example, with node #21 (which should be a **union**) shown as an **intersection**. (The symbols used in the diagrams, although distinct, are unreasonably similar.) With that correction, we reproduced all the paper's results.

John P. Mayberry,
Dept. of Mathematics
Brock University
St. Catharines, Ont
CANADA L2S 3A1

UNSOLVED PROBLEMS

Edited by: Richard Guy

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

A mod- n Ackermann Function, or What's So Special About 1969?

Jon Froemke and Jerrold W. Grossman

One of computer scientists' favorite functions is the *Ackermann function*, first studied by David Hilbert and Wilhelm Ackermann about 75 years ago [2]. It is recursive (i.e., computable), but it grows too fast to be primitive recursive (i.e., computable without using dirty tricks like double recursion or the operator "the least n such that"). A kind of inverse for this function (which grows excruciatingly slowly—it makes something like $\ln \ln n$ look like the U.S. national debt by comparison) enters into the efficiency analysis of some important algorithms, such as keeping track of the components of a graph as new edges are added.

If we restrict the range of the Ackermann function to a finite set (with a suitable "mod"-ification of its definition), then we might expect the exuberance of the original function to be reflected in rather chaotic behavior within this set. In fact we seem to find just the opposite, with the finitized Ackermann function petering out very quickly. We have many partial results about this mod- n Ackermann function, obtained using fairly straightforward ad hoc arguments as well as a little elementary number theory. We also have some intriguing experimental data. Perhaps some readers of this article can provide a more definitive description of what's going on.

To be specific, let \mathbf{N} denote the set $\{0, 1, 2, 3, \dots\}$ of natural numbers, and for each integer $n > 2$ let \mathbf{N}_n denote the set $\{0, 1, 2, \dots, n-1\}$ of natural numbers less than n . Define the **standard mod- n Ackermann function** from $\mathbf{N} \times \mathbf{N}_n$ to \mathbf{N}_n by

$$A_n(i, j) = \begin{cases} (j + 1) \bmod n & \text{if } i = 0 \\ A_n(i - 1, 1) & \text{if } i > 0 \text{ and } j = 0 \\ A_n(i - 1, A_n(i, j - 1)) & \text{if } i > 0 \text{ and } j > 0. \end{cases}$$

A (possibly) **nonstandard** mod- n Ackermann function is defined in the same way, except that the values $A_n(0, j)$ for $j = 0, 1, 2, \dots, n - 1$ are arbitrary. We will write A_n^s to refer specifically to the standard function. The value $A_n(i, j)$ is said to be in the i th column and j th row; we picture these values arranged as in Figures 1 and 2.

12	0	1	1	5	3	5	9	9	...
11	12	0	12	1	2	9	9	9	...
10	11	12	10	12	6	5	9	9	...
9	10	11	8	11	5	9	9	9	...
8	9	10	6	4	0	5	9	9	...
7	8	9	4	7	1	9	9	9	...
6	7	8	2	2	11	5	9	9	...
5	6	7	0	6	9	9	9	9	...
4	5	6	11	8	3	5	9	9	...
3	4	5	9	9	2	9	9	9	...
2	3	4	7	3	6	5	9	9	...
1	2	3	5	0	5	9	9	9	...
$j = 0$	1	2	3	5	0	5	9	9	...
<hr/>									
	$i = 0$	1	2	3	4	5	6	7	...

Figure 1. The standard mod-13 Ackermann function A_{13}^s .

6	6	3	0	0	0	0	...
5	0	4	0	4	0	4	...
4	3	0	0	0	0	0	...
3	5	5	0	4	0	4	...
2	6	3	0	0	0	0	...
1	0	4	0	4	0	4	...
$j = 0$	4	0	4	0	4	0	...
<hr/>							
	$i = 0$	1	2	3	4	5	...

Figure 2. A nonstandard mod-7 Ackermann function A_7 .

If we set $n = \infty$, then we obtain in the standard case one of the usual versions of the nonfinitized Ackermann function. It grows monotonically (and wildly) as i and j increase; for example, $A_\infty^s(2, 3) = 9$, $A_\infty^s(3, 3) = 61$, and $A_\infty^s(4, 3)$ has about 10^{20000} digits.

Let us adopt the following terminology. Denote the set of values that appear in the i th column by $P_n(i)$. Clearly $P_n(0) \supseteq P_n(1) \supseteq P_n(2) \supseteq \dots$; denote the intersection of this sequence, $\bigcap_{i=0}^\infty P_n(i)$, by P_n . If A_n becomes constant in some column i , i.e., $A_n(i, 0) = A_n(i, 1) = \dots = A_n(i, n - 1)$, then the function is said to have **stabilized** in column i (and clearly remains constant in all subsequent columns). The smallest i , if any, such that A_n has stabilized in column i is called the **stability number** of A_n , denoted by $s(n)$ in the case of the standard mod- n Ackermann function.

For $n < \infty$ only two kinds of asymptotic behavior are possible (since there are only finitely many different columns, and each column is uniquely determined by the one before it): either A_n stabilizes, or the columns are (nontrivially) **periodic**, i.e., for some $t > 1$, $A_n(i, j) = A_n(i + t, j)$ for all j and large enough i . In the nonstable case the smallest t for which this occurs is called the **period**.

Figures 1 and 2 illustrate the only two known ways in which any mod- n Ackermann function behaves asymptotically. In Figure 1 we see that $s(13) = 6$. This behavior, in which the function stabilizes fairly quickly, seems to happen in almost all cases, standard or not. On the other hand, in Figure 2, we see a nonstable situation for a nonstandard mod- n Ackermann function, in which the period is 2. It is easy to construct an example of this type for any even positive integer m , i.e., a nonstandard mod- n Ackermann function with period 2, whose columns eventually alternate between $(m, 0, 0, 0, \dots)$ and $(0, m, 0, m, \dots)$, in fact starting with a permutation of N_n in column 0 as long as $m > 2$.

Here is what we have found computationally. The only value of $n < 1,000,000$ for which the standard mod- n Ackermann function does not stabilize is $n = 1969$. (The first author's older child has been searching for some mystical significance to this property of his birth year.) For $n = 1969$ the period 2 behavior starts in column 8, with the columns alternating between $(1698, 0, 0, 0, \dots)$ and $(0, 1698, 0, 1698, \dots)$. For all other $n < 500,000$, the stability number for the standard function is at most 15, and is usually much less (for example, it often happens that $s(n) = 5$ and $P_n(5) = \{65533\}$). On the other hand, since $\lim_{n \rightarrow \infty} A_n^s(i, j) = A_\infty^s(i, j)$ for any fixed i and j , the function $s(n)$ is unbounded. We have also tried all possible starting columns for all $n \leq 10$, and there are no other patterns.

Here is some of what we know theoretically. First, P_n cannot be all of N_n ; in other words, at least some numbers have to disappear as we move from column to column. To prove this, suppose that $P_n = N_n$. Since $A_n(i + 1, 0) = A_n(i, 1)$, the number 1 cannot appear in column $i + 1$ except in row $n - 1$, or else $A_n(i + 1, 0)$ would be repeated. Hence 1 must appear in row $n - 1$ in every column from 1 on. But the only way that $A_n(i + 2, n - 1)$ gets to be 1 is for $A_n(i + 2, n - 2)$ to be $n - 1$ (because 1 appears only in row $n - 1$ of column $i + 1$). Hence $n - 1$ must appear in row $n - 2$ in every column from 2 on. Similarly, $n - 2$ must appear in row $n - 3$ in every column from 3 on. Eventually this says that 2 must appear in row 1 in every column from $n - 1$ on, which is absurd, since if 2 appears in row 1 in column i , then it appears in row 0 in column $i + 1$. The "line-'em-up" argument used in this proof seems useful in deriving other results as well.

Once we know at least that $P_n \neq N_n$, under what conditions can we go the whole distance and prove that $|P_n| = 1$ (i.e., A_n stabilizes)? On the one hand, we can prove that $|P_n| = 1$ if $0 \notin P_n$ or $1 \in P_n$. Our strongest result is that the standard mod- n Ackermann function stabilizes if n has a prime factor p such that $2^{j+3} \equiv 3 \pmod{p}$ has no solutions; this is the case for $p = 2, 3, 7, 17, 31, 41$ and 43 , to name the first few. From still another perspective, we can show that the two situations discussed above (and illustrated in Figures 1 and 2) are the only possible asymptotic behaviors when $|P_n| \leq 4$ or the period is 2. Open questions abound, such as whether 1969 is the only counterexample to stability in the standard case, or how to compute $s(n)$ efficiently.

As a final variation, we can run the Ackermann function "in reverse" to generate for each n a canonical but random-looking permutation of $N_n - \{1\}$, somewhat in the spirit of the shuffles reported on by David Gale [1]. Again we start with $A(0, j) = j + 1$ for all $j > 0$, but we set $A(0, 0) = 0$. The procedure for producing column $i + 1$ from column i is as follows: $A(i + 1, 1) = A(i, 0)$, and for $j \neq 1$, $A(i + 1, j) = A(i, k + 1)$, where $A(i, k) = j$. The first few columns are shown in Figure 3.

Note that each column can be obtained from the column *following* it by applying our original construction. It is easy to show that this function is well-

\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
8	9	9	9	9	9	9	9	9	...
7	8	8	8	8	8	8	8	3	...
6	7	7	7	7	7	7	5	0	...
5	6	6	6	6	6	0	0	8	...
4	5	5	5	5	3	5	6	2	...
3	4	4	4	2	0	6	2	4	...
2	3	3	0	0	5	4	4	6	...
1	2	0	2	3	4	2	3	7	...
$j = 0$	0	2	3	4	2	3	7	5	...
	$i = 0$	1	2	3	4	5	6	7	...

Figure 3. The Ackermann function in reverse.

defined in each column; it gives a permutation of $\mathbb{N} - \{1\}$ that leaves $A(i, j) = j + 1$ for all $j > i$. Here one might ask, for example, whether every positive integer $j \neq 1$ appears infinitely often in each row other than row j . As of yet, we have no answers.

REFERENCES

1. D. Gale, Mathematical entertainments, *The Mathematical Intelligencer*, 14, no. 1 (1992) 54–57.
2. J. W. Grossman and R. S. Zeitman, An inherently iterative computation of Ackermann’s function, *Theoretical Computer Science*, 57 (1988) 327–330.

Department of Mathematical Sciences
Oakland University
Rochester, MI 48309-4401
grossman@vela.acs.oakland.edu

Erratum: Contrary to the information we received some months ago and which was published in the October issue of this MONTHLY, we have just been advised that Professor Emeritus G. H. Hunt is alive and well: our deepest apologies and our very best wishes to him for a long life ahead.

—*American Mathematical Monthly* 75, (1968) p. 1145.

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before July 31, 1993 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10282. *Proposed by Paul Erdős, Hungarian Academy of Sciences, Budapest, Hungary.*

Let A, B, C be the vertices of a triangle inscribed in a unit circle, and let P be a point in the interior of the triangle ABC . Show that

$$|PA| \cdot |PB| \cdot |PC| < \frac{32}{27}.$$

10283. *Proposed by Feng Luo, University of California, San Diego, CA, and Richard Stong, University of California, Los Angeles, CA.*

Let D be a convex polygonal region in the plane and let f be a bounded convex (and hence continuous) function on the interior of D .

(a) Show that f extends to a continuous function on all of D .

(b) Show that the analogous result does not hold if D is the unit disk.

10284. Proposed by Liang-shin Hahn, University of New Mexico, Albuquerque, NM.

For each positive integer l , show that there exists a positive integer n and a partition of $\{1, \dots, n\}$ as a disjoint union of two sets A and B , such that for $1 \leq i \leq l$,

$$\sum_{a \in A} a^i = \sum_{b \in B} b^i.$$

10285. Proposed by Frank Schmidt, Arlington, VA.

Let e_n , respectively o_n , denote the number of unlabeled graphs on n vertices having an even, respectively odd, number of edges. Show that $e_n \geq o_n$ for all n .

10286. Proposed by Călin Popescu, Université Catholique de Louvain, Louvain-La-Neuve, Belgium.

Let $a_0 + a_1x + \dots + a_mx^m$ be a polynomial with real coefficients and $(-1)^ma_m > 0$. Suppose that all roots of this polynomial are positive real numbers less than 1. Prove that

$$(-1)^{p+1} \sum_{k=n}^{m-p-1} (-1)^k \binom{k}{n} \sum_{h=0}^m a_h \sum_{i+j=k} (-1)^j \binom{h}{i} \binom{m-h}{j} > 0$$

for all nonnegative integers n and p whose sum is less than m .

10287. Proposed by Dr. A. Keith Austin, The University of Sheffield, Sheffield, England.

We have a doubly-infinite (i.e. indexed by \mathbb{Z}) row of squares and we start with counters in those squares to the left of some point (e.g. those with negative index). For a fixed positive integer k , the allowable moves consist of selecting k consecutive squares, discarding one of the counters in those squares, and rearranging the remaining counters within the k selected squares (with at most one counter in a square). Prove or disprove that there is an integer $N = N(k)$ such that no sequence of moves will allow a counter to be placed N squares into the region which originally contained no counters.

10288. Proposed by Bruce R. Johnson, University of Victoria, Victoria, B. C., Canada.

From an urn containing b balls, numbered from 1 to b , balls are drawn one at a time with replacement until the accumulated sum of all numbers drawn is at least equal to a positive integer n . Let X_n denote the amount by which the accumulated sum exceeds n . Find $\lim \mathbf{E}(X_n)$ or show that this limit does not exist.

10289. Proposed by David M. Bloom, Brooklyn College of CUNY, Brooklyn, NY.

For $x > 1$, consider the inequality

$$a\sqrt{x} + (1-a)\left(\frac{x+1}{2}\right) < e^{-1}x^{x/(x-1)}.$$

(a) If $a \geq 1/3$, show that the inequality holds for all $x > 1$.

(b) If $a < 1/3$, show that there is some $x > 1$ for which the inequality is false.

NOTES

(10286) In the innermost sum, we employ the convention that a binomial coefficient $\binom{h}{i}$ is zero unless $0 \leq i \leq h$. (10287) For a similar result in the traditional peg solitaire, see John D. Beasley, *The Ins & Outs of Peg Solitaire*, Oxford University Press, 1985, chapter 12. (10288) The number of balls, b , is fixed so its value may appear in the answer. For each n , X_n is a random variable whose expectation is denoted by $E(X_n)$. (10289) The case $a = 1/2$ appeared as problem 1365 in *Mathematics Magazine*.

SOLUTIONS

Scrambling Points on the Unit Circle

6647 [1991, 63]. *Proposed by Andrew Vince, University of Florida, Gainesville, FL.*

Let $S_n = \{1, \zeta, \zeta^2, \dots, \zeta^{n-1}\}$ be the set of n -th roots of unity and suppose f is any function on S_n into the set of complex numbers of absolute value one. For every positive integer k less than $n/2$ prove that there exist integers i and j such that

$$|\zeta^i - \zeta^j| \geq |1 - \zeta^k| \geq |f(\zeta^i) - f(\zeta^j)|.$$

Note: The proposer had been confused with a different A. Vince in the original presentation of the problem.

The solution will be divided into two parts, called lemmas 1 and 2. In these proofs, distance will be the shortest distance measured along the unit circle in units of one- n -th of a circle, so that all the distances between points in S_n are integers. Since Euclidean distance is a monotonic function of this distance, the desired results can be obtained from the corresponding results for this distance. Also, the set of all images of points in S_n under f will be referred to as $f(S_n)$.

Lemma 1. *There exists a closed arc of the unit circle of length k which contains at least $k + 1$ points of $f(S_n)$.*

Solution 1 by O. P. Lossers, Eindhoven University of Technology, Eindhoven, The Netherlands. Let the elements of $a_j \in f(S_n)$ be ordered by their arguments

$$0 \leq a_0 \leq \dots \leq a_{n-1} < n$$

and extend the indexing to all $j \in \mathbb{Z}$ so that $a_{j+n} = a_j$. Then,

$$(a_k - a_0) + \dots + (a_{nk} - a_{(n-1)k}) = nk$$

so at least one of the terms, say $(a_{lk} - a_{(l-1)k})$, has a value not exceeding the average of the n terms which is k . Now, the $k + 1$ points a_j with $(l - 1)k \leq j \leq lk$ all lie in the closed arc from $a_{(l-1)k}$ to a_{lk} , which was chosen to have length at most k .

Solution II by Sharad Kanetkar, University of Massachusetts, Boston, MA. Consider n closed arcs on the unit circle, each of length k , chosen so that their beginning points are equally spaced (at distance 1) and so that at least one such beginning point coincides with one of the points of $f(S_n)$.

Each point of the circle lies in at least k such arcs, and the endpoints lie in $k + 1$ arcs. In particular, each point of $f(S_n)$ lies in at least k arcs and at least one of them lies in $k + 1$ arcs. Thus the total number of incidences of points of $f(S_n)$ with these arcs is at least $nk + 1$. For Lemma 1 to be false, however, each of these arcs would contain at most k elements of $f(S_n)$ and the total number of incidents would be at most nk .

Lemma 2. *If $A \subset S_n$ has $k + 1$ elements, then there is a pair of elements in A whose distance is at least k .*

Solution by the Editors, based on an idea of Sharad Kanetkar. For each element $a_i \in A$, there is a set $A_i \subset S_n$ consisting of a_i together with the $k - 1$ consecutive elements of S_n clockwise from a_i and the $k - 1$ consecutive elements of S_n counterclockwise from a_i .

If, for any i , some element a of A lies outside A_i , then a_i and a are the desired elements. Otherwise, A lies wholly within each A_i . By DeMorgan's laws, this is equivalent to saying that the union of complements C_i of the A_i is a subset of the complement of A in S_n . However, the complement of A contains $n - k - 1$ elements and we shall show that the union of the C_i must contain at least $n - k + 1$ elements.

To prove the latter claim, note that each A_i contains $2k - 1$ consecutive points of S_n , so that C_i contains $n - 2k + 1$ consecutive points of S_n . Start from a point not in the union of the C_i (if no such point exists, the claim is clearly true) and look at the C_i in clockwise order starting from this point. The first C_i gives us $n - 2k + 1$ points and each of the k subsequent C_i gives at least one point beyond (in the clockwise sense) the previous C_i since the C_i are distinct intervals.

Editorial comment. The result clearly follows from the lemmas: Lemma 1 gives a set A of $k + 1$ of the ζ^i satisfying the condition on the $f(\zeta^i)$ and Lemma 2 allows a pair of these elements to be selected to satisfy

$$|\zeta^i - \zeta^j| \geq |1 - \zeta^k|$$

as well.

All successful solvers except the proposer followed the outline presented here, although Lemma 2 seemed rather elusive.

Lemma 1 was obtained also by L. E. Mattics and R. Stong. The proposer's solution and one other were judged to be unconvincing.

Another Block-Walking Identity

E 3439 [1991, 437]. *Proposed by Jane Friedman, Widener University, Chester, PA.*

If M and N are nonnegative integers, prove that

$$\binom{M+N}{M} = \sum_{0 \leq a \leq (M-1)/2} \binom{M-a-1}{a} \binom{N+a}{2a+1} + \sum_{0 \leq a \leq M/2} \binom{M-a}{a} \binom{N+a}{2a}.$$

Solution I by Kiran S. Kedlaya (student), Georgetown Day High School, Washington, DC. Consider walks on the coordinate plane from $(0, 0)$ to (M, N) , where each step increases either the x coordinate or the y coordinate by 1. The number of such walks is $\binom{M+N}{N}$. Now consider the set of points

$$\mathbf{S} = \{(M, 0), (M-1, 0), (M-2, 1), (M-3, 1), \dots, \\ \times (M-2a, a), (M-2a-1, a), \dots\}.$$

Every walk meets \mathbf{S} and so has a unique last point of intersection with \mathbf{S} . We count the walks by this last intersection with \mathbf{S} .

First we count the walks whose last intersection with \mathbf{S} is at $(M-2a, a)$. There are $\binom{M-a}{a}$ ways to reach that point and $\binom{N+a}{2a}$ ways from there to (M, N) , so there are $\binom{M-a}{a} \binom{N+a}{2a}$ walks in this class.

Next we count the walks whose last intersection with \mathbf{S} is at $(M-2a-1, a)$. There are $\binom{M-a-1}{a}$ ways to reach that point, and the next step must be to $(M-2a-1, a+1)$. From there, we can reach (M, N) in $\binom{N+a}{2a+1}$ ways, so there are $\binom{M-a-1}{a} \binom{N+a}{2a+1}$ walks in this class.

Composite solution II by Rolf Richberg, RWTH Aachen, Aachen, Germany and Chris Wildhagen, Rotterdam, The Netherlands. Since $\binom{M+N}{N}$ is the coefficient of $x^M y^N$ in $(1-x-y)^{-1}$, it suffices to show that $R(x, y) = \sum_{M, N \geq 0} r_{M, N} x^M y^N = (1-x-y)^{-1}$, where $r_{M, N}$ is the sum on the right, which extends without change to all positive a . Interchange the order of summation, so

$$R(x, y) = \sum_{a \geq 0} x^{2a} y^a \sum_{M \geq 0} \binom{M-a}{a} x^{M-2a} \sum_{N \geq 0} \binom{N+a}{2a} y^{N-a} \\ + \sum_{a \geq 0} x^{2a+1} y^{a+1} \sum_{M \geq 0} \binom{M-a-1}{a} x^{M-2a-1} \sum_{N \geq 0} \binom{N+a}{2a+1} y^{N-a-1}.$$

Since $\sum_{r \geq 0} \binom{r+j}{j} z^r = (1-z)^{-j-1}$, we have

$$R(x, y) = \sum_{a \geq 0} \frac{x^{2a} y^{2a}}{(1-x)^{a+1} (1-y)^{2a+1}} + \sum_{a \geq 0} \frac{x^{2a+1} y^{a+1}}{(1-x)^{a+1} (1-y)^{2a+2}} \\ = \frac{1 + \frac{xy}{1-y}}{(1-x)(1-y)} \sum_{a \geq 0} \left(\frac{x^2 y}{(1-x)(1-y)^2} \right)^a \\ = \frac{1-y+xy}{(1-x)(1-y)^2 - x^2 y} = \frac{1}{1-x-y}.$$

Hence $r_{M,N}$ is indeed the coefficient of $x^M y^N$ in $(1 - x - y)^{-1}$.

Editorial comments. The method of Solution I displays a combinatorial elegance, but would be difficult to find without having both the given form of the sum on the right and the simple closed form $\binom{M+N}{N}$ on the left, while the method of Solution II provides a systematic way of simplifying the sum. Paul Deiermann took yet another approach, reducing the identity to the Vandermonde identity $\sum_{j=0}^M \binom{M}{j} \binom{N}{j} = \binom{M+N}{N}$ (see Dean S. Clark, "On Some Abstract Properties of Binomial Coefficients", this MONTHLY, 89 (1982), 433–443).

Solved also by J. Balogh (student, Hungary), K. L. Bernstein, D. Callan, R. J. Chapman (U. K.), P. Deiermann, L. Denenberg, M. Dindos (Czechoslovakia), J. S. Frame, W. P. Gerlach, F. T. Howard, N. Komanda, J. H. van Lint (The Netherlands), O. P. Lossers (The Netherlands), R. Martin (student), S. G. Penrice, J. Sarkar, A. J. Stam (The Netherlands), Anchorage Math Solutions Group, and the proposer.

Subspaces of the Space of Endomorphisms

E 3444 [1991, 438]. *Proposed by Wilbur Jónsson, McGill University, Montreal, Canada and Gérard Letac, Université Paul Sabatier, Toulouse, France.*

Let E be a finite-dimensional vector space over a field K and let $L(E)$ be the space of endomorphisms of E . Suppose L_1 and L_2 are subspaces of $L(E)$ such that $L(E) = L_1 + L_2$ and $xy + yx = 0$ for all (x, y) in $L_1 \times L_2$.

- (i) If $\text{char}(K) \neq 2$, prove that either $L_1 = 0$ or $L_2 = 0$.
- (ii) If $\text{char}(K) = 2$, show that the conclusion of (i) need not hold.

Solution by Steve Ott, Lexington Community College, Lexington, KY. If $\dim(V) = n$, then $L(E)$ can be represented by the set of all n by n matrices with entries in K . Put $I = B + C$, with $B \in L_1$ and $C \in L_2$, where I is the identity matrix. For any $A \in L_1$, we have $AB + BA = A(I - C) + (I - C)A = 2A - (AC + CA) = 2A$. In particular, $2B^2 = 2B$, which implies $B^2 = B$ since $\text{char}(K) \neq 2$. Hence multiplying $AB + BA = 2A$ by B on the left gives $BAB + BA = 2BA$, or $BAB = BA$, and multiplying by B on the right gives $BA + BAB = 2AB$, or $BAB = AB$. Thus $BA = AB$. Now we have $2A = AB + BA = 2AB$, or $AB = A = BA$. Finally $I = B + C$ implies $A = AB + AC = A + AC$, and thus $AC = 0$. Similarly, $CA = 0$. By applying the same argument to an arbitrary $D \in L_2$, we obtain $CD = D = DC$ and $BD = 0 = DB$ for all $D \in L_2$.

Now consider an arbitrary $X \in L(E)$, with $X = X_1 + X_2$ where $X_1 \in L_1$ and $X_2 \in L_2$. Using the results above, we have $BX = BX_1 + BX_2 = X_1 = X_1B + X_2B = XB$, and B commutes with every n by n matrix. It is well known (by induction on n , for example) that this implies B is a multiple of the identity. Hence $B = bI$ and $C = (1 - b)I$.

If $b = 0$, we now have $A = AB = 0$ for all $A \in L_1$, or $L_1 = \{0\}$. If $b \neq 0$, then our results for arbitrary $D \in L_2$ imply $0 = BC = b(1 - b)I$, which implies $b = 1$ and $C = 0$. Hence $D = DC = 0$ for all $D \in L_2$, and $L_2 = \{0\}$.

For part (ii), let K be the field of integers mod 2 and suppose $\dim(E) = 2$, so $L(E)$ is the set of 2 by 2 binary matrices. Let $L_1 = \{0, I\}$ and $L_2 = L(E)$. Then L_1 and L_2 are subspaces of $L(E)$ with $L_1 + L_2 = L(E)$ and $XY + YX = 0$ for all $X \in L_1, Y \in L_2$, but neither L_1 nor L_2 is the trivial space.

Editorial comment. Several solvers expressed their work entirely in the language of unital associative algebras. The property of $L(E)$ used was either *simplicity* (no two-sided ideals) or *connectedness* (no idempotents other than 0 and 1). F. J. Flanigan also investigated examples for rings without identity.

Solved also by M. Barr (Canada), K. Benbury, D. Callan, R. J. Chapman (U. K.), G. Ehrlich, M. Falkowitz (Israel), F. Flanigan, P. Freyd, W. H. Gustafson, L. Hogben, N. Komanda, C. Lanski, O. P. Lossers (The Netherlands), A. Nijenhuis, National Security Agency Problems Group, and the proposers.

A Riemann Sum Revealed

E 3446 [1991, 552]. *Proposed by Jean-Pierre Grivaux, Lycée Chaptal, Paris, France.*

Suppose $0 \leq u_0 < 1$ and $0 < \lambda < 1$. Define a sequence $\{u_n\}_{n=0}^\infty$ by putting

$$u_{n+1} = u_n + \lambda \sqrt{1 - u_n^2} \quad (n = 0, 1, 2, \dots).$$

For small n the sequence $\{u_n\}$ is real and increasing. Let $N = N(\lambda, u_0)$ be the first integer for which $u_N \geq 1$. For fixed u_0 find

$$\lim_{\lambda \rightarrow 0+} \lambda N(\lambda, u_0).$$

Solution by Kenneth F. Andersen, University of Alberta, Edmonton, Alberta, Canada. The value of the limit is $\cos^{-1} u_0$. Observe first that

$$0 < u_{n+1} - u_n \leq \lambda, \quad (n = 0, 1, 2, \dots, N-1).$$

In particular, $u_N - u_{N-1} \leq \lambda$ and since $u_{N-1} < 1 \leq u_N$, we have $\lim_{\lambda \rightarrow 0} u_{N-1} = 1$. Thus, $u_0 < u_1 < \dots < u_{N-1} < 1$ is a partition of $[u_0, 1]$ with mesh size at most λ so that

$$\begin{aligned} \cos^{-1} u_0 &= \int_{u_0}^1 \frac{du}{\sqrt{1-u^2}} = \lim_{\lambda \rightarrow 0} \left(\sum_{n=0}^{N-2} \frac{u_{n+1} - u_n}{\sqrt{1-u_n^2}} + \frac{1 - u_{N-1}}{\sqrt{1-u_{N-1}^2}} \right) \\ &= \lim_{\lambda \rightarrow 0} \sum_{n=0}^{N-2} \lambda + \lim_{\lambda \rightarrow 0} \sqrt{\frac{1 - u_{N-1}}{1 + u_{N-1}}} \\ &= \lim_{\lambda \rightarrow 0} \lambda(N-1) \\ &= \lim_{\lambda \rightarrow 0} \lambda N. \end{aligned}$$

Editorial comment. Although most solvers used the approach through Riemann sums, it is also possible to obtain properties of the u_n directly. In particular, the recurrence arises in the solution of $u' = \sqrt{1-u^2}$ by Euler's method with step size λ . Since this problem deals with expressions which are only well-behaved for $|u| < 1$, technical difficulties arise in attempting to use a continuous process to model the behavior as $\lambda \rightarrow 0$.

Solved also by U. Abel (Germany), R. J. Chapman (U. K.), M. Falkowitz (Israel), E. A. Grove & V. Lj. Kocic & G. Ladas, P. G. Kirmser, N. Komanda, O. P. Lossers (The Netherlands), H. Morris, F. C. Rembis, E. Suárez (Spain), D. Velleman, Y. Yildirim, National Security Agency Problems Group, and the proposer.

The Determinant of a Hankel Matrix

E 3447 [1991, 553]. *Proposed by Peter Borwein, Dalhousie University, Nova Scotia, Canada.*

If p and q are relatively prime positive integers with $0 < p < q$, let $T_{p,q}$ be the q by q matrix in which the element in the i th row and j th column is

$$t^{\lfloor p(i+j-1)/q \rfloor} \quad (i, j = 1, 2, \dots, q).$$

Show that

$$\det T_{p,q} = (-1)^{\lfloor q/2 \rfloor} t^{(p-1)(q-1)+p} (t-1)^{q-1}.$$

Here $\lfloor \cdot \rfloor$ denotes the greatest integer (or “floor”) function.

Solution by O. P. Lossers, Eindhoven University of Technology, Eindhoven, The Netherlands. We shall give a proof by determining the factors of the determinant. The magnitude of every term in the permutation expansion of $\det T_{p,q}$ is t raised to a power of the form $\sum_{i=1}^q \lfloor (i + \sigma(i) - 1)p/q \rfloor$, where σ is a permutation of $\{1, \dots, q\}$. For every σ , we have

$$\sum_{i=1}^q \left\lfloor i \frac{p}{q} \right\rfloor + \sum_{i=1}^q \left\lfloor (\sigma(i) - 1) \frac{p}{q} \right\rfloor \leq \sum_{i=1}^q \left\lfloor (i + \sigma(i) - 1) \frac{p}{q} \right\rfloor \leq pq.$$

The inequality on the right is sharp if and only if $\sigma(i) = q + 1 - i$ for all i , and the leftmost expression equals $\sum_{i=1}^q \lfloor ip/q \rfloor + \sum_{i=1}^q \lfloor (j-1)p/q \rfloor = 2\sum_{i=1}^{q-1} \lfloor ip/q \rfloor + p$.

The sum $\sum_{i=1}^q \lfloor ip/q \rfloor$ counts the lattice points in the interior of the triangle whose corners are $\{(0,0), (q,0), (q,p)\}$ (grouped by columns). Since p, q are relatively prime, there are no lattice points on the diagonal except $(0,0)$ and (p,q) , so by symmetry we have $2\sum_{i=1}^{q-1} \lfloor ip/q \rfloor = (p-1)(q-1)$. Altogether, these results imply that every contribution to $\det T_{p,q}$ is divisible by $t^{(p-1)(q-1)+p}$ and that $\det T_{p,q}$ is a polynomial in t with leading term $(-1)^{\lfloor q/2 \rfloor} t^{pq}$.

It thus suffices to show that $\det T_{p,q}$ is divisible by $t^{(p-1)(q-1)+p}(t-1)^{q-1}$, which is a polynomial in t of degree pq . We already have the first factor. Since every element of the matrix is a power of t , and $\det T_{p,q}$ is not identically zero, we may subtract the first row from all other rows to obtain a matrix in which $q-1$ rows are multiples of $t-1$. Hence we have $q-1$ factors of $t-1$ in the computation of the determinant, which completes the proof.

Editorial comment. Robin J. Chapman and José Heber Nieto (independently) obtain the same result by exhibiting a sequence of elementary row operations giving the desired factors.

Solved also by J. T. Bruening, R. J. Chapman (U. K.), Th. Honold (Germany), K. S. Kedlaya (student), J. H. Nieto (Venezuela), I. A. Sakmar (Turkey), and the proposer.

A curious property of $1/7$

6661 [1991, 559]. *Proposed by Jeffrey C. Lagarias, AT&T Bell Laboratories, Murray Hill, NJ, and Thomas Zaslavsky, SUNY, Binghamton, NY.*

A curious property of $\frac{1}{7}$ is that to two decimal places it equals $.02 \times 7$. Add $.02^2 \times 7$ and you obtain $\frac{1}{7}$ to four decimal places. Add $.02^3 \times 7$ and you obtain it

to six places (with an error of 1 in the last place), and so on. In fact, $\frac{1}{7}$ equals 7 times the sum of a geometric series whose ratio has a terminating decimal expansion:

$$\frac{1}{7} = 7 \times \sum_{i=1}^{\infty} (.02)^i.$$

Which positive integers N have a similar representation,

$$\frac{1}{N} = N \sum_{i=1}^{\infty} r^i,$$

where r is a terminating decimal?

Solution by Kevin Ford (student), University of Illinois, Urbana, IL. Only the integers 1, 2, 3 and 7 have such a representation. If the integer N has such a representation, then $r = 1/(N^2 + 1)$ and hence $N^2 + 1 = 2^c 5^d$, where c and d are non-negative integers and $0 \leq c \leq 1$ since $N^2 \not\equiv 3 \pmod{4}$. If $d \leq 4$, the only solutions are those stated. We suppose then that $d > 4$ and break the argument into two cases, according to whether $c = 0$ or $c = 1$. In both cases we will be using the facts that $\mathbf{Z}[i]$ is a unique factorization domain, and that $1 \pm i$ and $2 \pm i$ are primes in $\mathbf{Z}[i]$. Throughout (x, y) denotes the greatest common divisor of x and y , u denotes a unit in $\mathbf{Z}[i]$ (so that $u \in \{\pm 1, \pm i\}$), and $\lfloor x \rfloor$ denotes the greatest integer $\leq x$ for real x . All variables except u represent rational integers.

Case I ($c = 0$). First note that d is odd, for otherwise we would have an integer solution to $x^2 + 1 = y^2$ with $|y| > 1$. The equation $N^2 + 1 = 5^d$ factors as

$$(N + i)(N - i) = (2 + i)^d (2 - i)^d.$$

Since $(N + i, N - i) = (N + i, 2i) = ((2 + i)(2 - i), 2i) = 1$, we have either

$$\left\{ \begin{array}{l} N + i = u(2 + i)^d \\ N - i = u^{-1}(2 - i)^d \end{array} \right\} \quad \text{or} \quad \left\{ \begin{array}{l} N + i = -u^{-1}(2 - i)^d \\ N - i = -u(2 + i)^d \end{array} \right\},$$

both of which imply $2i = u(2 + i)^d - u^{-1}(2 - i)^d$. Let $v_d = (2 + i)^d + (2 - i)^d$ and $w_d = i((2 + i)^d - (2 - i)^d)$. Both v_d and w_d are real, hence either $v_d = \pm 2$ or $w_d = \pm 2$. But $v_0 = 2$, $v_1 = 4$ and $v_{k+2} = 4v_{k+1} - 5v_k$. Thus $4|v_d$ if d is odd, and we are left with

$$w_d = (-1)^{(d-1)/2} \cdot 2 \left\{ -1 + 2^2 \binom{d}{2} - 2^4 \binom{d}{4} + 2^6 \binom{d}{6} - \cdots \right\} = \pm 2.$$

Both sides must agree modulo 8, so

$$\binom{d}{2} - 2^2 \binom{d}{4} + 2^4 \binom{d}{6} - \cdots = 0.$$

Define l , e_k and f_k by $2^l \parallel (d-1)$, $2^{e_k} \parallel k$ and $2^{f_k} \parallel 2^{2k-2} \binom{d}{2k}$. From the relation

$$\binom{d}{2k} = \frac{d(d-1)}{2k(2k-1)} \binom{d-2}{2k-2},$$

it follows that $f_k \geq 2k - 2 + l - (1 + e_k) \geq l - 3 + 2k - \lfloor \log k / \log 2 \rfloor \geq l$ if $k \geq 2$. This implies $\binom{d}{2} = d(d-1)/2 \equiv 0 \pmod{2^l}$, which is impossible.

Case II ($c = 1$). Clearly N is odd, so write $N = 2s + 1$. We then have $s^2 + (s + 1)^2 = 5^d$, which factors as $(s + si + i)(s - si - i) = (2 + i)^d(2 - i)^d$. Also

$$\begin{aligned}(s + si + i, s - si - i) &= (2s, s + si + i) = (2, s + si + i) \\ &= (2, (2 + i)(2 - i)) = 1,\end{aligned}$$

so either

$$\left\{ \begin{array}{l} s + si + i = u(2 + i)^d \\ s - si - i = u^{-1}(2 - i)^d \end{array} \right\} \quad \text{or} \quad \left\{ \begin{array}{l} s + si + i = u(2 - i)^d \\ s - si - i = u^{-1}(2 + i)^d \end{array} \right\}.$$

Since $(1 + i)(s + si + i) + (1 - i)(s - si - i) = -2$, we have

$$-2 = u(1 + i)(2 + i)^d + u^{-1}(1 - i)(2 - i)^d,$$

which implies $2i^e = (1 + i)(2 + i)^d \pm (1 - i)(2 - i)^d$ where $0 \leq e \leq 3$. Let $a_d = (1 + i)(2 + i)^d + (1 - i)(2 - i)^d$ and $b_d = i((1 + i)(2 + i)^d - (1 - i)(2 - i)^d)$. Both a_d and b_d are real, so either $a_d = \pm 2$ or $b_d = \pm 2$. Expanded using the binomial theorem, a_d and b_d each have the form

$$2 \left(\pm 1 \pm 2 \binom{d}{1} \pm 2^2 \binom{d}{2} \pm 2^3 \binom{d}{3} \pm \cdots \right), \quad (1)$$

where the sequence of signs (\pm) on the left has period 4 and depends on the residue class of d modulo 4. For $0 \leq k \leq 3$, let A_k (resp. B_k) be the equation $a_d = \pm 2$ (resp. $b_d = \pm 2$) obtained when $d \equiv k \pmod{4}$. The value of the right-hand side of each equation ($+2$ or -2) is uniquely determined by reducing the equation modulo 8. The following table lists the first four signs on the left of (1) and the value of the right side for each equation:

Equation	Sequence of Signs	Right side
A_0	$++--$	2
A_1	$-++-$	2
A_2	$--++$	-2
A_3	$+-+-$	-2
B_0	$-++-$	-2
B_1	$--++$	2
B_2	$+-+-$	2
B_3	$++--$	-2

Four of these equations do not hold modulo 32. Using the fact that $\binom{d}{3}$ is odd if and only if $d \equiv 3 \pmod{4}$, equation B_1 reduces to $d^2 - 2d - 1 \equiv 0 \pmod{8}$, equation B_2 reduces to $-d^2 \equiv 0 \pmod{8}$, equation B_3 reduces to $-d^2 + 2d + 5 \equiv 0 \pmod{8}$, and equation A_3 reduces to $5 - d^2 \equiv 0 \pmod{8}$. The falsehood of each of these congruences is easily verified.

In each of the remaining equations A_k and B_k (with $k = 0, 1$ or 2), we first introduce l and f_h such that $2^l \parallel (d - k)$ and $2^{f_h} \parallel 2^h \binom{d}{h}$. If $k = 0$ or $k = 2$ we also introduce g_k such that $2^{g_h} \parallel h(h - 1)(h - 2)$ and use the relation

$$\binom{d}{h} = \frac{d(d - 1)(d - 2)}{h(h - 1)(h - 2)} \binom{d - 3}{h - 3}$$

to obtain

$$f_h \geq h + l + 1 - g_h \geq h + l + 1 - \left(1 + \left\lfloor \frac{\log h}{\log 2} \right\rfloor\right) \geq l + 4 \quad (h \geq 6).$$

Also $f_5 \geq l + 4$, hence modulo 2^{l+5} the left-hand side of each equation reduces to the first five terms. If $k = 1$, the relation $\binom{d}{h} = d(d-1)/h(h-1)\binom{d-2}{h-2}$ similarly leads to $f_h \geq l + 4$ for $h \geq 5$. Multiply each equation by 3 and reduce modulo 2^{l+5} . Equation A_0 reduces to $4d(d-2)[-3 - 2(d-1) + (d-1)(d-3)] \equiv 0$, equation B_0 reduces to $4d[d - 2(d-1)(d-2) - (d-1)(d-2)(d-3)] \equiv 0$, equation A_1 reduces to $4(d-1)[3 + 3d - 2d(d-2) - d(d-2)(d-3)] \equiv 0$, and equation A_2 reduces to $4d(d-2)[3 + 2(d-1) - (d-1)(d-3)] \equiv 0$. Since none of the bracketed expressions are divisible by 4, none of these congruences can hold.

Editorial comment. Several readers pointed out that this problem appeared as E 2511 [1975, 73; 1976, 291]. The published solution at that time consisted mainly of a reference to L. J. Mordell, *Diophantine Equations*, Academic Press, 1966. However, equations of this type appear to be easier to solve than to research (see T. Nagell, "The Diophantine equation $x^2 + 7 = 2^n$ ", *Ark. Math.*, 4 (1961), 185–187). The solution above shows that consideration of binomial coefficients modulo powers of 2 suffice to complete a proof based on factorization in the Gaussian integers. Another approach, taken by the proposers and a majority of readers, begins by using the parity of c and d to reduce to the equation $x^2 - Dy^2 = -1$ for $D = 1, 2, 5$ or 10 with the extra condition that no prime other than 2 or 5 can divide y . The theory of the Pell equation gives expressions for y in terms of the fundamental unit of the field generated by \sqrt{D} . In each case, the powers of 2 and 5 dividing these expressions can be determined. Except in the cases given, the expressions are easily shown to be larger as real numbers than the largest $2^a 5^b$ which divides them. The proposers point out that this approach allows the effective determination of the finite set of solutions of every equation of the form

$$N^2 + 1 = p_1^{a_1} \dots p_k^{a_k}.$$

Solved also by H. L. Abbott & H. I. Freedman (Canada), D. Callan, R. J. Chapman (U.K.), E. Curtin, J. Drost, C. Friesen (Canada), S. M. Gagola Jr., C. P. Grant, J. Grantham (student), S. Hahn, R. Holzager, I. Kastanas, P. Lindstrom, J. Mann, J. Rickert, R. M. Robinson, L. Simons, N. Singer, P. G. Walsh (student, Canada), and the proposers. Six incorrect or incomplete solutions were received.

The Herglotz Trick Applies

E 3454 [1991, 646]. *Proposed by John A. Baker, University of Waterloo, Ontario Canada, and motivated by the work of Hans G. Kelloerer, Munich, Germany.*

Suppose $f: \mathbb{R} \rightarrow \mathbb{R}$ is a function such that

$$2f(x+1) = f(x) + f(2x)$$

for all real x .

Prove that if f is twice differentiable, then f must be constant.

Solution 1 by Klaus Zacharias, Berlin, Germany. Taking $x = 0$ and $x = 1$ gives

$$f(0) = f(1) = f(2). \quad (1)$$

We differentiate

$$2f(x+1) = f(x) + f(2x)$$

and obtain

$$2f'(x+1) = f'(x) + 2f'(2x),$$

$$2f''(x+1) = f''(x) + 4f''(2x).$$

So the continuous function $F = f''$ satisfies the functional equation

$$2F(x+1) = F(x) + 4F(2x).$$

Putting $2x = y$ gives

$$F(y) = \frac{1}{2}F\left(\frac{y}{2} + 1\right) - \frac{1}{4}F\left(\frac{y}{2}\right). \quad (2)$$

Let $a \geq 2$ and $I = [-a, a]$, and put

$$M = \max_{y \in I} |F(y)|.$$

If $y \in I$, then obviously $(y/2) \in I$ and $(y/2) + 1 \in I$. From (2) we get

$$|F(y)| \leq \frac{1}{2}M + \frac{1}{4}M = \frac{3}{4}M \quad \text{for all } y \in I$$

and hence $M \leq .75M$. Thus $M = 0$. Because $a \geq 2$ was arbitrary, we have $F(y) = 0$ everywhere. From $f'' = F = 0$ follows

$$f(y) = Ay + B$$

and, because of (1), we obtain that $f(x)$ is constant.

Solution II by David Cruz-Urbe, SFO, student, University of California, Berkeley, CA. We will show that f is constant on the interval $[0, 2\pi]$. An identical argument applies to all intervals $[2n\pi, 2(n+1)\pi]$, $n \in \mathbb{Z}$.

Since $f \in C^2$ on $(0, 2\pi)$, it is of bounded variation. Thus its Fourier series converges uniformly on every closed interval contained in $(0, 2\pi)$ (see Yitzhak Katznelson, *An Introduction to Harmonic Analysis*, Dover, 1976, p. 53). Therefore, on any interval on which it makes sense, the given identity becomes

$$2 \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{in\pi x} = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx} + \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{2inx}.$$

We can equate coefficients between sides. If k is odd we get

$$2\hat{f}(k)e^{ik} = \hat{f}(k).$$

Hence $(2e^{ik} - 1)\hat{f}(k) = 0$, which implies that $\hat{f}(k) = 0$ since $|2e^{ik} - 1| > 1$. By induction on r , the same argument shows that $\hat{f}(2^r k) = 0$ for all $r \in \mathbb{N}$. Since k was an arbitrary odd number, this gives $\hat{f}(n) = 0$ for all $n \in \mathbb{Z}$ with $n \neq 0$. Thus $f(x) = \hat{f}(0)$, a constant.

Editorial comment. The idea of solution I is the ‘‘Herglotz trick.’’ See R. Remmert, *Theory of Complex Functions*, Springer, 1991, and S. Bochner, ‘‘Review of Herglotz’ Gesammelte Schrfiten’’, *Bull. Amer. Math. Soc.* 1 (1979), 1021. Variants on this solution were submitted by most solvers. Adam Riese noted that the continuity of f'' is used only to show that f'' is bounded on each closed interval. Charles Voas proved that, for $\rho > 1$, a C^1 function g satisfying

$$g(\rho x) = \alpha g(x) + \beta g(x+1)$$

for all $x \in \mathbb{R}$ with complex constants α and β satisfying

$$|\alpha + \beta| < \min(1, \rho - 2|\beta|)$$

must be identically zero. This applies to the problem at hand by taking $g = f'$.

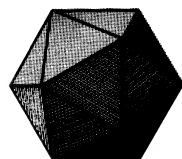
Solved also by S. E. Arteaga, R. J. Chapman (U. K.), M. Dindos (Czechoslovakia), R. High, R. Holzager, I. Kastanas, O. P. Lossers (The Netherlands), A. Riese, R. M. Robinson, M. Roth & O. Šuch (Canada), A. Swett, B. J. Vekatachala (India), D. Velleman, C. Voas, and the proposer. One incorrect solution was received.

Collaborating editors: *David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttmann, Frank B. Miles, Richard Pfiefer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, Edward T. H. Wang, and William E. Watkins.*

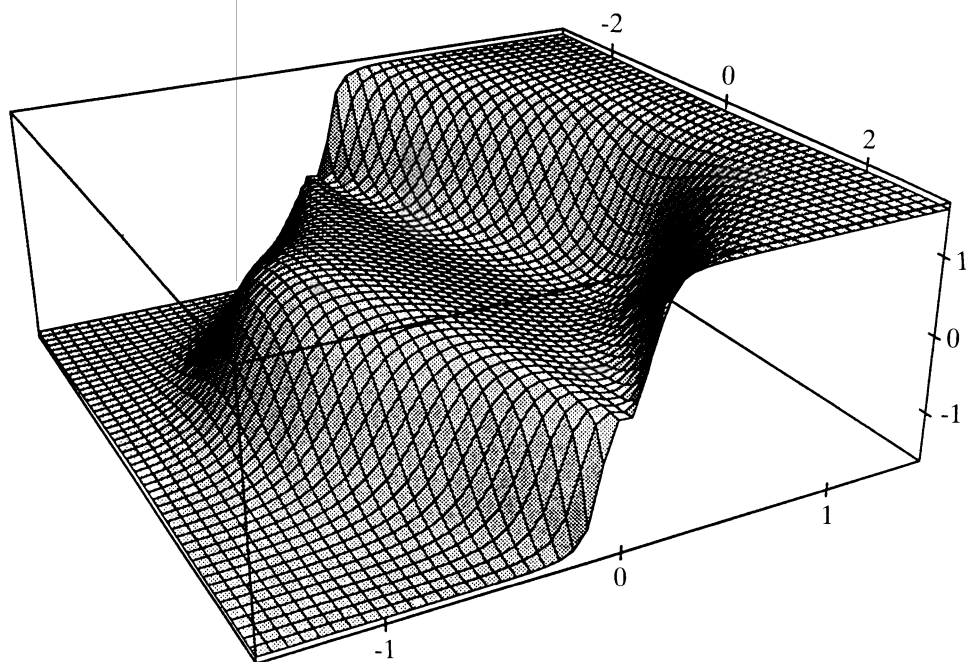
Answer to Picture Puzzle:
(on page 175.)

They *are* the same man: D. H. Lehmer, taken about 1939 and 1975.

The American Mathematical Monthly



Volume 100, Number 3 / MARCH 1993



AN OFFICIAL PUBLICATION OF THE MATHEMATICAL ASSOCIATION OF AMERICA

NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTELEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

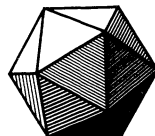
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International, Serials coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 3 / MARCH 1993
(ISSN 0002-9890)



Contents

ARTICLES

Szeged in 1934 / EDGAR R. LORCH 219

Tarski's High School Identities / STANLEY BURRIS
and SIMON LEE 231

Aperiodic Chaotic Orbits / STEVEN N. MACEACHERN
and L. MARK BERLINER 237

Bricklaying and the Hermite Normal Form / WILLIAM J. GILBERT 242

A Characterization of Inner Product Spaces / NEIL FALKNER 246

Polar Area Is the Average of Strip Areas / GILBERT STRANG 250

Counting Critical Points of Real Polynomials in Two Variables /
ALAN DURFEE, NATHAN KRONENFELD, HEIDI MUNSON,
JEFF ROY, and INA WESTBY 255

FEATURES

COMMENTS 218

NOTES 272

PICTURE PUZZLE 282

THE AUTHORS 283

LETTERS 285

UNSOLVED PROBLEMS 287

Parker's Permutation Problem Involves the Catalan Numbers 287

PROBLEMS AND SOLUTIONS 290

REVIEWS

Ethnomathematics: A Multicultural View of Mathematical Ideas
by Marcia Ascher / JUDITH V. GRABINER 304

TELEGRAPHIC REVIEWS 309

Szeged in 1934

Edgar R. Lorch
(edited by Reuben Hersh)

This article is based on two manuscripts written by Edgar R. Lorch for a proposed book on mathematics in Hungary, to be edited by Reuben Hersh. The book project was abandoned, for lack of manuscripts. After Lorch's death, his manuscripts were combined and edited by Reuben Hersh to produce this article.

1. WHY SZEGED? The year 1933 was dreary and dismal for most people. This was precisely at the end of my studies for the doctorate. By some miracle, I was spared unemployment. I was awarded a National Research Council Fellowship, for a year of postdoctoral study at Harvard University under the guidance of Marshall Stone.

There were four N.R.C. fellows at Harvard: Magnus Hestenes, David Nathan, Deane Montgomery, and me. Montgomery and I formed a seminar, to lecture each other on the material in Oswald Veblen's *Analysis Situs*. We were joined by Norman Steenrod, who shone by his determination to find out what was really going on.

These fellowships had been severely cut, both in number and in stipend. Nevertheless, even with the reduced stipend of \$1,600 for twelve months, I managed to live in Boston like a Bohemian, dividing my activities between wooing the recalcitrant muse of mathematics and indulging in the follies of youth: drinking beer, going to symphony concerts, and jogging in the park. This extra year at Harvard was supposed to give us a "coat of varnish," as one of my friends put it. Whether it turned us into gentlemen or scholars is a moot question. It did provide a line in my curriculum vitae. Future employers were impressed.

Stone told me that the two mathematicians in all the world who could be most helpful to my development were John von Neumann and Frederick Riesz. (His Hungarian name, Frigyes, became Frederick when anglicized). von Neumann's name was well known to me, of course. Of Riesz, situated at the other end of the world, I was only vaguely aware. But Stone's remark was prophetic.

Stone and I agreed that my energies were best spent digging deeper into Hilbert and Banach spaces. That meant a frontal attack on Stone's Colloquium Series Publication, *Linear Transformations in Hilbert Space*, a remarkable 622-page book more often quoted than read. Stone's style has sometimes been subject to comment. It is infinitely correct but, like many of his other qualities, carried out "à outrance," with some bizarre results. For example, one of his theorems (p. 590) takes two complete pages—for the statement. The last chapter of the book fills 218 pages.

Stone had a special sense of humor. At one of our infrequent meetings I mentioned some problems as possible candidates for research topics. About the best one of my problems, Stone said, "Oh, I don't know. Somebody must have

worked on that problem already.” The following week, while browsing in Widener Library, I came upon an article containing the complete solution to the problem. The author: M. H. Stone.

All of us Fellows were terrified what would happen to us if we couldn’t locate a spot for next year. The only offer I heard about was to Deane Montgomery: a \$1,600 assistant professorship at the University of Nevada. I applied for an extension of my National Research Fellowship. At this time the political super-potentates of the mathematical scene were centered in Princeton, N.J., where the Institute for Advanced Study had recently been established. The School of Mathematics was its leading school. There were about five mathematics professors at the Institute. In order to further distinguish them from ordinary mortals teaching at Columbia, Yale, or other universities, each professor had an assistant. There was tremendous variation in the duties of these assistants. It was traditional belief that Einstein’s assistant did nothing. The only requirement for him was to be a Jewish exile from Nazi Germany. Hermann Weyl’s assistant had normal duties: preparing in mimeographed form his professor’s lectures on group theory. I cannot imagine what Alexander’s or Veblen’s assistants did—probably not much.

In early spring these potentates got together, counted up the mathematical plums to be handed out for the year, and made a list of the available talent who constituted the target space on which these plums were to be mapped. Then they sent the customary letters to the candidates: a three-paragraph personal letter to the candidates who had been hit, and a one-paragraph note of non-success to the poor souls who did not make it. One day I learned that one of my friends had received his letter—a good one. I gingerly went home, and sure enough, there was a letter from the Institute. The type-print covered the whole page—success! I would be able to live another year.



Edgar R. Lorch, Szeged, 1934

The letter was really exciting. I was being offered the job of assistant to John von Neumann! I had heard him lecture several times. He was brilliant, spoke very fast, his English was quite fluent, he made remarkably few errors. A characteristic one was to talk about “infinite serious” for infinite series. No one ventured to correct his few lapses. I had met him recently at a party. The high point of the evening was a recitation race between him and Norbert Wiener. Somehow, someone recited a line from Lewis Carroll’s “The Hunting of the Snark.” Norbert,

with his usual ebullience and sonorous voice, began reciting from line 1. Johnny started off in pursuit. Norbert accelerated, but Johnny came up even. We held our breaths as the lines poured out, on and on until they reached the end in a dead heat.

I made a trip to Princeton and met with Veblen, who was then running the Institute. "What," asked I, "are the duties of an assistant to Professor von Neumann?" Veblen answered with a mixture of surprise and disdain, that a mere private second class should ask such a question about a four star general. His answer staggered me. Here were the four principal duties of von Neumann's assistant:

1. To attend von Neumann's lectures on operator theory on Mondays, Tuesdays, and Wednesdays, take copious notes, complete unfinished proofs, see them through the secretarial jungle, and promptly circulate them to all American university libraries. This task alone was consuming the entire energies of a younger person, who had to be not only well-meaning but sharp, fast, clever, and tough. These notes ran to over 600 pages.
2. To be von Neumann's assistant as Editor of the *Annals of Mathematics*. This meant reading through every manuscript accepted for publication, underlining Greek letters in red and German letters in green, and circling italics. Also writing in the margins all necessary instructions to printers. The following anecdote illustrates the hazards of being editorial assistant of the *Annals* in the early thirties. A manuscript was submitted by the brilliant Soviet mathematician, Lev Pontryagin. Since paper was then exceptionally scarce in the Soviet Union, Pontryagin had taken wrapping paper, torn it into appropriate-sized pieces, and gone to work on his typewriter. Unfortunately, Pontryagin was blind. The wrapping paper was torn unevenly, and a good portion of the words and symbols in the margins were missing. No matter. The *Annals* editorial assistant retyped the paper, supplying all the missing symbols. What a hero!
3. To go once a week to the printers of the *Annals of Mathematics* in Baltimore in order to instruct them in the art of typesetting mathematical symbols with subscripts, superscripts, subsubscripts, etc. The *Annals of Mathematics* had been printed in Hamburg, Germany by the firm of Lütke and Wolk. In view of increasing anti-Semitism under Hitler, the German connection was given up in favor of printing in the United States. But no American printer had ever before set up mathematical symbols! They were complete illiterates. Solution: let von Neumann's assistant teach them!
4. To translate into English von Neumann's numerous 100-page papers. Now that von Neumann was a professor in an American institute, it was thought that his papers should appear in English, not German. Since von Neumann was provided with an assistant, it was natural that the assistant should do this.

Items three and four were on the table for the first time. The first two had somehow been handled by the previous assistant. I left my meeting with Veblen in a downcast mood. Here I had the opportunity to work next to the most brilliant mathematician of his generation. But the job entailed such onerous duties that only someone with an iron constitution could survive. My constitution, it so happened, was not made of iron. It was made of reeds and bamboo sticks, very satisfactory under moderate pressures, but completely incapable of standing a huge overload. But what choice had I?

A few days after returning to Cambridge, I received a letter from the Dean of the Graduate School at my alma mater, Columbia University. The contents made me radiant. In view of my “outstanding work in scholarship and research,” the letter said, I had been awarded a Cutting Traveling Fellowship for the following academic year. There were no conditions on the award, except to use it for travel, presumably in Europe. The stipend was \$1,800 (more than I was currently receiving as a National Research Council Fellow). What should I do? I debated with myself and discussed it with others. I already had in mind the professor whom I wanted to visit: Frederick Riesz of Hungary—the man suggested by Professor Stone. Riesz worked in a town I had never heard of, Szeged. I went to my atlas and found it: the second city in Hungary. There was a university: *Ferencz-Jozsef Tudomány Egyetem*. I was mouthing my first words in a strange new language. If John von Neumann was the acknowledged genius of modern mathematics, Frederick Riesz was the dean of functional analysts. He was not well known to the world at large, but the cognoscenti had the highest respect for him. In the field that interested me, Riesz was the classic leader.



Professor F. Riesz Szeged, 1934

I wrote to Riesz asking if I could spend some months with him. I wrote to von Neumann and Veblen, explaining that my nervous constitution would not allow me to perform adequately the duties of Von Neumann’s assistant. (I heard subsequently that the position that had been offered to me was divided into four pieces. Heaven knows, each quarter was substantial enough.) Riesz wrote me a short, pleasant letter accepting my proposal. The die was cast. Off to Szeged!

II. SZEGED. In the early twentieth century, there were only a few universities in Hungary. The one in Budapest, the nation’s capital, was the most important. Next was the one in Kolozsvár, capital of Transylvannia, a rich region of which all Hungarians were proud. In fact, Kolozsvár was Hungary’s second capital, the seat of a host of administrative offices. When Frederick Riesz was a young man, he had a friend and rival in a closely related field, Lipót Fejér. Both were outstanding analysts, both at the University of Kolozsvár. In retrospect, Riesz was much deeper and was to have much greater influence internationally. When a position opened

at Budapest, it was Fejér who received the call. This meant that Riesz remained at Kolozsvár.

The dismemberment of the Austro-Hungarian Empire by the Treaty of Trianon following World War I was a catastrophe for Hungary, economically and psychologically. It lost two thirds of its territory. All of Transylvania was awarded to Romania. The entire Hungarian administrative apparatus at Kolozsvár (henceforth to be known as Cluj) was moved into the new, smaller Hungary. This included the university. Where should it go?

On the southern border of the new Hungary was a sleepy town called Szeged, now the second largest in the country. This city was the natural bastion against the Yugoslavs and Romanians, and the center of a rich agricultural area called the *Nagyalföld*. Here the university was transported, with its entire faculty—including F. Riesz. The mathematics faculty founded a journal, the *Acta Scientiarum Mathematicarum*. (For short it is called the Szeged *Acta*, to distinguish it from the original *Acta Mathematica* in Sweden and the *Acta Mathematica Hungarica* in Budapest). Within a few years the Szeged *Acta* had a world-wide reputation. Every serious mathematics library receives it. So, Szeged is now known all over the world—at least to mathematicians.

In 1934 Szeged was an agricultural town of 120,000. There may have been three automobiles. They belonged to the mayor and the chief of police. There must have been a taxi also, for I remember riding in it. The town lived on three activities. It was a garrison town, swarming with officers and soldiers. The hated Romanians and Yugoslavs were only fifteen kilometers away. It was a market for the huge fertile plain surrounding it. And it was a university town, with full range of studies, distinguished faculty, and tens of hundreds of serious students. From the American point of view, it was at the end of the world.

Placid is the word which best fits Szeged. *Unhurried* is appropriate too. With broad streets (one or two would qualify as avenues) and low, separated buildings, it was a city one could live in—providing one had something to do. It was not a city for excitement. There wasn't one tourist in town. In 1934, there was one foreigner. Virtually no one from Budapest had ever been in Szeged. This is probably true even today. Very few streets were paved. A few were cobble-stones. The rest were packed dirt. Sidewalks existed in the innermost part. Elsewhere, one walked in the street, stepping aside for the farmer's cart plodding along or the horse-drawn coach, its driver lashing at the none too elegant horse.

The deepest first impression was the horses, and the carriages they pulled. The horses, horses, horses. And in summer, the flies. These horses would invariably salute the bystander as they went by. One heard the clop-clop of their hooves as they approached from behind. At the moment of passing, just as the coachman was cracking an oath and his whip, the nags would greet you. Up went their tails, out came five or six steamy, greenish-brown balls which splattered the pavement. It was like the salute of a platoon of soldiers passing a reviewing stand. Oh horses of Szeged, where are you now? You deserve to be in the Valhalla for quadrupeds!

Every day at twelve o'clock, Szeged observed the great social event of the day: the *korzo*. The entire town gathered, or so it seemed, on one of the few fashionable streets, some three or four blocks long, and paraded on foot back and forth for thirty to forty minutes, looking right and left with a view to recognizing the greatest number of people and greeting the greatest number of friends. Everyone dressed well, and if possible better. In the winter, one put on one's finest overcoat. Every man wore a hat. The greeting, to a man or to a lady, was raising the hat completely off the head, simultaneously making a pronounced bow, all the

while continuing to walk briskly forward. This courteous greeting was a Hungarian custom ingrained firmly in youth, and not easily forgotten in later years. I remember receiving such a greeting in 1950 in Cambridge, Massachusetts, across a 70 meter avenue, from that most courteous genius, John von Neumann. Sometimes the two greeters would stop to converse a minute. Then the words exchanged were “*Szervusz*” between two young men, and “*Kezét csókolom*” from a man to a lady. This *korzo* performance kept up furiously for thirty to forty minutes. The crowd was so thick that one felt one was in Times Square. Then all of a sudden the crowd would dissolve, the street became deserted, and Szeged returned to its usual somnolence.

Postdoctoral students lived in dormitories such as the Eötvös Kollégium. Each had an adequate room, with cot, table and hard chair. The room was kept in order by a hall boy, who would run errands for a tip. The entrance to the Kollégium was locked at 11 p.m. Access for latecomers was by ringing a bell to waken the concierge. After midnight you had to pay a fee. Poor students couldn’t afford it, and were in their rooms by 10 p.m. The “rich” could spend their evenings drinking wine in the cafes and return at all hours. This was especially true during carnival, when merriment went on until the wee hours of the morning.

The carnival period was the welcome sign that the long, bitter, sunless winter was coming to an end, and the dreary routine of work and study was going to be broken. The merriment took the form of good-natured drinking and frenzied dancing. Each café had one or two orchestras, and the music continued with scarcely a pause until after four in the morning. Scores of young men were present, including students and officers. Also present was a group of modest young ladies, each accompanied by her chaperone. The boys outnumbered the girls about four to one. The girls wore formal evening gowns. But remember, we are in Szeged, not Budapest, and the year is 1935, not a very affluent time. So the dresses were of satin or cretonne, definitely not too sparkling. During the course of the carnival a girl might be seen alternating between two dresses.

The favorite music was, of course, the Viennese waltz. It was played with a frenetic élan from beginning to end, passing through all the dances of Johann Strauss and Ferenc Lehár. At the end of each dance, the orchestra would immediately break into the next swirl, even before the vigorous applause for the last dance had died out.

The boys formed groups of four or five who would dance with the same girl. As soon as one boy had exhausted himself, having spun his partner and himself to the point of dizziness, his companion would come up, tap him on the shoulder and shove off with the hapless girl. At the end of that dance, a third boy would come up and repeat the act. The young lady had started the evening like a cool, fresh flower, but little by little she had become a steamy mass of limp flesh, flushed in the face and perspiring under the arms. And let her beware of begging off, claiming that she was tired! Thenceforth she would be ostracized, and not dance again. It was pitiful to see each girl pirouetted around the floor while her squadron of predators sat at a table, sipping wine a bit too freely, and from time to time addressing a courteous phrase to the chaperone, who was having thoughts about their qualifications for matrimony, and about how times had changed (for the worse) since she was young.

Still, with this merriment, there was political tension in the air. It was all over Europe, and Szeged shared it. Newspapers were avidly read; some people spent two hours a day on them. The situation with the immediate neighbors, Yugoslavia and Romania, was quite dangerous. An innocent visit by taxi to their triple border



Edgar R. Lorch, in the library at Szeged, 1934

with Hungary, just a few kilometers away, evoked an alarming response from the border guards on the other side, enough to make one cut short the visit. But everyone who had perspective had his eye on Germany. The last war had been bad enough. What would happen now? On a visit to Budapest I met the Minister of Commerce and his charming, well-to-do family, including his beautiful twenty-year-old daughter. The father stated his problem bluntly. He was a Jew. He could read the writing on the wall. He saw what was going to happen in Germany. He would be the happiest man in the world if I would marry his daughter and take her away from Europe. Poor man! What happened to him? And what happened to the sweet young girl?

III. RIESZ FRIGYES. At a European congress of mathematicians around 1910, three or four outstanding young mathematicians sipping tea in a café decided to send a postcard to a highly esteemed English colleague, G. H. Hardy. No one signed his name. Instead, each put down the one formula he had discovered which had made him famous. Riesz put down his representation formula for the general linear functional on the space of continuous functions $C_{(0,1)}$, which is now known as the Riesz representation theorem: $Ff = \int_0^1 f(x) d\mu(x)$. Needless to say, Hardy needed no prompting to unravel the card. Among the cognoscenti Frederick Riesz was already a world-known figure.

Besides his representation theorem, his most famous contributions include: the central role in creating the theory of compact linear operators; recreating the Lebesgue integral without relying on measure theory; the use of subharmonic functions as basic tools in potential theory; introducing the spaces L^p , H^p , and C into mathematics, and the basic work on their linear functionals; the ergodic theorem; the proof that monotonic functions are differentiable almost everywhere; the Riesz-Fischer theorem, which is a central result about abstract Hilbert space, and is also essential for the proof of the equivalence between Schrödinger's wave mechanics and Heisenberg's matrix mechanics.

Riesz was a quiet man who sometimes may have given the impression that he was unapproachable. That impression was incorrect. In his later life in Szeged, he acted the part of a "vieux garçon," always pleasant with the people around him, usually deep in thought. People there could hardly have conceived that he was a mathematical genius. Local people (excluding university intimates) were amazed

that someone had come all the way from New York to be close to him. Many times I was taken aside and asked, "Is he really that famous?"

His lack of attention to others may have given the appearance of displeasure. He knew his own worth fully, and calmly steered his course. Undoubtedly he was troubled at having been kept in Kolozsvár and Szeged most of his life, not receiving recognition from his country by a call to Budapest until he was late in years.

The people he came in contact with—excluding, of course, his mathematician colleagues—thought he was a strange, inoffensive man. To them he was a sort of teddy-bear—short, rather corpulent, unhurried in movement, slow, mumbly, and parsimonious in his speech. He could invariably be found, either in his easy chair at the Institute, at lunch at his table in the Hotel Tisza, or at supper in his club. His meals would last three hours—well, perhaps not always, but certainly two and a half. After eating copiously with his napkin under his chin, he would light up his cigar, have the waiter bring the day's newspapers, and plunge into a complete reading of at least five of them. From time to time a cloud of smoke would rise from behind his paper, or he would throw back his head, close his eyes, and after a pause mutter "*Ja uj*, I see now." Then back to the paper. Finally, he would rise and slowly, very slowly, walk to his apartment.

During this walk, he would frequently stop and turn to me, his eyes blazing with pleasure. He would get close, probably because he was short-sighted, and push me back with his stomach. Then he would tell me, ever so briefly, about a new idea for a proof made "*pour épater le bourgeois*." He loved mathematical pranks. For example, in the fields of measure, integration, and differentiation, the classic order of development is the one I have just written down. Riesz showed that the three subjects can be developed in any order. He loved being a bad boy, upsetting preconceived ideas.

Riesz and I both had quiet personalities. When we walked together, each of us was absorbed in his own mathematical thoughts. But I could interrupt him with a question, which he accepted without fuss, to be disposed of in any of a variety of ways. To: "Do you know an example of such and such"; or "Why in the theory of absolutely convergent series does it seem that . . .," he might answer: "Aha, that is a good question." After an abundant meal and his reading of all the local newspapers, he would tap me on the arm and say, "I have your example. Consider . . ." Or simply, "Oh, no. You must not ask that question," and half wink. Once, after I suggested a problem, he repeated it slowly twice to make certain he understood. Then, absolute silence. After a half hour, he exclaimed, "*Ja uj*. I have it. Yes, I have it," and he leaned to me, his eyes full of joy and mischief. But that was the end of it. He never made me privy to his solution.

Riesz's younger brother, Marcel, was also a very distinguished mathematician, a brain export from Hungary to Sweden. Their interests were not identical, but they did publish one very important result together, known to all as the theorem of F. and M. Riesz. Fred was very fond of his brother, and spent every summer with him. The two were a prize exhibit at any mathematical congress. However, Marcel knew what the score was. Once at a meeting, a mathematician seeing them exclaimed, "Ah, here are the two Riesz!" "No," returned Marcel, "there is only one Riesz."

Riesz was a dangerous man to collaborate with. He was constantly having new ideas, and his latest brain child was his favorite. This could have disconcerting results for his collaborator. The experience of his former assistant, Tibor Radó, illustrates this.

During the academic year, Riesz lectured on measure theory and functional analysis. Radó would take copious notes. When summer came Riesz would leave for a cooler spot (Győr). Radó would sweat it out in Szeged for three months, writing up the lecture notes for publication in the fall. At the end of September, Riesz would put in his first day at the Institute. Radó would come to the library to greet his superior, proudly carrying a stack of eight hundred hand-written pages, which he placed in Riesz's lap with a look of great satisfaction. Riesz would glance at the bundle, and raise his eyes with a mixture of kindness and thankfulness, and at the same time with a spark of merriment, as if he had pulled off a fast one. "Oh, very good, very good. Yes, this is very nice. But I tell you, during the summer I had an idea. We will do it another way. You will see when I give the course. You will like it!" This took place many years in a row. The book was not written until eighteen years later, with Béla Szőkefalvi-Nagy as co-author.

Riesz was a quiet, law-abiding citizen. He respected the laws which mesh us in. But he also had a sense of humor, and enjoyed enormously any situation which made the "administration" look foolish. I was involved in one of those situations. When I arrived in Hungary, I had a visa valid for three months. Around the end of November, I received notice from the police that my visa was about to expire. Riesz forthwith asked his assistant (secretaries were unknown at the university) to prepare a statement that I was carrying out research under his supervision, and requesting my visa be extended for six months. On Friday afternoon the letter was typed on Institute stationery, and signed by Riesz with all his titles. The next morning, equipped with the letter and my passport, I went to the police. The day was beautiful, lighted but not heated by a late fall sun. The city was quiet and seemed well-disposed. After a short wait I was shown to the Chief. He received me courteously, and read the letter. But then he gave signs of indecision. He called in his second in command. They began an animated discussion. At first the second in command just listened. Then he seemed to become completely of the opinion of his Chief. The two of them came over to their petitioner, who was getting nervous, and began a long explanation in Hungarian, which at first I did not understand. By and by, I got the point. The letter was fine, except that the Institute's rubber stamp had not been added to it. Without the rubber stamp, the letter was not valid.

I rushed to the Institute. It was Saturday, and no one was in. But I had a key. I thought I would simply go in, go to the one and only desk, open its one and only drawer, take out the one and only rubber stamp, stamp it on my letter, and then rush back to the police with my letter in A-1 order.

Easier said than done! When I opened the drawer, I found not one rubber stamp, but fifteen: Airmail, First Class, Registered, Printed Matter, and so on. They were all Greek to me. I decided on the following solution. (To this day I don't know where I got the nerve). I took a blank piece of Institute stationery, and on it printed each of the fifteen rubber stamps. Then I ran back to the police, jubilantly pointed at the fifteen imprints on the paper, and said to them, "Just show me which one you want, and I'll go back to the Institute and put it on the letter for you!" The police didn't understand one word of this speech, but they grasped the situation: they were dealing with either a child, a feeble-minded person, or a dangerous character. And what kind of offices are they running at the University, with sacred seals left within reach of any American amateur spy? After a quick consultation they gave me my visa and ushered me out of the office, with severe admonitions which were not understood.

On Monday, I reported these happenings to Riesz. When he understood my offense against officialdom, he collapsed in laughter. For weeks he repeated the

story to everyone. I am sure his opinion of me shot up. It was the kind of thing he would gladly have done himself in younger days.

When I first met Riesz, I expected to communicate in French. Although he still wrote French of considerable elegance, he was no longer in his French period. I never heard him utter more than a passing phrase in German—at most, he might tell a German-Hungarian joke. He had been studying English for the past three years, and so English became the language between us. His study of English had been exceptionally thorough. Not only had he assiduously studied grammar and composition, he had read much literature, including a good number of the plays of G. B. Shaw. As a distinguished professor, it was understood that he was capable of learning anything he wanted to in the seclusion of his room. In fact, he never had an English teacher with whom to exchange a single word. (There was one English teacher in Szeged, an unmarried lady who had been living there for almost twenty years, teaching the few adventurous souls who were intrigued by things British.)

It was to me that Riesz uttered his first words of English, after three years of study. The results were strange and at times, with the best of good will on both sides, slow. Sometimes Riesz corrected my English pronunciation. For example, once I asked whether a goose that had just been brought was well-cooked. He did not understand until it came to him in two syllables. “Coo-kud,” he corrected, and set to with a twinkle of satisfaction.

Spoken languages are a common topic of conversation among Hungarians, since their native tongue is so impenetrable to foreigners. In one group the question was once raised, what languages did Riesz speak? “Riesz speaks no language,” answered a waggish tongue. This was in a way a compliment. Riesz’ mind was so much faster than his tongue that he had pretty much given up trying to get his tongue to keep up with his mind. He was such a perfectionist that at the moment he was starting to send a message, he would already have improved it twice, so he just suppressed the whole thing. But of the written word he was a master.

During the four or five holidays and long weekends during the year, Riesz always asked me to join him on an excursion. Our home in Budapest was the Hotel Gellért, at the foot of the rocky cliffs facing the Danube. It was probably the finest hotel in the city, known for relaxed lounges, beautiful rooms, delicious cuisine and, of course, its thermal baths. We spent many hours in its pool. I still see next to me the walrus-like body of F. Fiesz, swim-floating very slowly while dreaming up some theorem. In the afternoon we would visit local colleagues at Fejér’s home and hear the latest international gossip. Once during the late fall we took a three-kilometer walk in the Tátra Mountains of north-east Hungary. The walk took less than three hours—but not much less.

Riesz and I collaborated on one joint paper while I stayed with him. I did not have the experience so poignantly described by Radó. Our paper concerned the formula $A = \int dE(\lambda)$ for unbounded self-adjoint transformations in Hilbert space. The question, of course, was to go from the bounded case (known for thirty years) to the unbounded. Von Neumann and Stone had accomplished this by using bilinear forms and Lebesgue-Stieltjes integration, from which the final result could be pieced together, providing one was still following. Riesz had looked at this problem, and was not satisfied with the existing proofs. We joined forces to produce two new proofs clearing away the complexities of integration theory, showing that the unbounded case can be reduced to a sequence of bounded cases which can then be pieced together. (I went back to this formula for the last time in 1950, showing it can be obtained by complex integrals of the Cauchy type, all in the uniform topology of operators.) The joint paper was written up by me and sent to

Ray's Visit to Szeged in May 1989

The purpose of our visit to Hungary in May 1989 was to celebrate a collaboration between American and Hungarian mathematics of old days (1934). The arrangements fell mainly in the hands of Hungarian Italianists who of course invited me as their American colleague to a lecture in their department. The intermediary between us and the Italian Department in Szeged was Péter Sárközi, Chair of Hungarian Literature at the University of Rome who had organized step by step in detail our stay in Hungary.

The "homecoming" celebrated at the József Attila Tudományegyetem by Ray's old mathematical friends was a touching event. I sat with a group of young students in the back of the classroom. While my husband was filling a wide, beautifully clean blackboard with white mathematical formulas, three old gentlemen in the front row were nodding approvingly. The three of them (all tall, straight, elegantly dressed) questioned and answered afterwards, laughing heartily at some detail of the past and quizzing each other on the development of a famous theorem. Finally, they posed for pictures in the mathematical library, and in front of the "József Attila" building. Everything around us looked like those three men: elegant, decorous, serene.

Our days in Szeged glided by peacefully, sunny and serene: meetings at the University, lunches at the Hotel Tisza, dinners at the "Hungarian," where gypsy violins played so passionately that I felt like jumping up and dancing the csárdás. We lived in a Collegium called "Herman Ottó," a ten-story prefabricated building, a student dormitory, reasonably solid, comfortable and very clean, and at walking distance from the river Tisza, which destroyed the town almost completely in 1887. The river has a personality of its own: wide and overflowing with water, it has a menace which keeps the town on edge. Ray felt completely at home. He was peacefully working at his mathematics.

Maristella de Panizza Lorch, Professor of Italian
Columbia University

the *Transactions* of the A.M.S. in September 1935, after I had returned to New York. Riesz never had a chance to have a better idea on how to do it.

While I lived in Szeged, I published a paper on functions of self-adjoint transformations in which I replaced the previous definition by bilinear forms and Lebesgue-Stieltjes integration with a theory of measure determined by a resolution of the identity, where the measure of a set is a closed linear manifold. This measure has the virtues that the measure of the intersection and union of sets is the intersection and union of the measures, and the measure of a set essentially determines the set. Riesz was much interested in this paper, and made many useful suggestions while I was writing it.

The optimal relation between a mentor and his disciple is seldom achieved. The relation between Riesz and me was optimal. It was warm, intimate, continuous, without pressure, very calm, leaving each of us free to develop his own thoughts. My stay with him was a perfect amalgam of a healthy, pleasurable life and an

uninterrupted communication of mathematical ideas. Access to mutual discussion was free, but never overused. Frederick Riesz was indeed a perfect teacher and a warm companion.

At the end of May I left Szeged. I was grateful for all he had done, but it was only much later that I appreciated reasonably well his contribution to my education. On the overnight train from Budapest to Venice, a young Hungarian man accompanied by a young woman kindly asked me in poor German to leave them the compartment for the night. They had just been married that day and were beginning their honeymoon. On arrival at Venice the next morning, he opened the corridor window and got off the train, while his bride, who spoke only Hungarian, lowered their suitcase through the window. Then a terrible thing happened. The train pulled out, with her still on it. The poor girl was cut off from everything that she knew, everything she had been living with. I knew what she felt. Hadn't I too just been cut off from my previous life?

REFERENCES

1. E. R. Lorch, Functions of self-adjoint transformations in Hilbert space, *Acta Sci. Math. Szeged*, 7 (1934), 136–146.
2. F. Riesz and E. R. Lorch, The integral representation of unbounded self-adjoint transformations in Hilbert space, *Trans. Amer. Math. Soc.*, 39 (1936), 331–340.
3. F. Riesz and B. Szökefalvi-Nagy, *Leçons d'analyse fonctionnelle*, Akademiai Kiado, Budapest, 1952.
4. M. H. Stone, *Linear transformations in Hilbert space and their applications to analysis*, Amer. Math. Soc. Colloquium Pub., 15, New York, 1932.
5. O. Veblen, *Analysis Situs*, 2nd edition, American Mathematical Society, New York, 1931.

Mathematics Department
University of New Mexico
Albuquerque, NM 87131

THE CHAUVENET PRIZE

The Committee on the second award of the Chauvenet Prize recommended that the award be made to Professor T. H. Hildebrandt of the University of Michigan for his paper on "The Borel theorem and its generalizations" published in the *Bulletin* of the American Mathematical Society, volume 32(1926), pages 423–474. Professors A. J. Kempner and D. R. Curtiss for the committee stated that "the paper presents in clear and elementary fashion, and with adequate references to literature, the development of a rather broad range of ideas and results connected with the Borel theorem. It gives the reader a compact picture not easily obtained by independent reading of the scattered literature on the subject." The Trustees adopted the recommendation and the award was announced at the annual business meeting. The prize, \$100 in cash, has since been conferred.

—*American Mathematical Monthly*
 37, (1930) p. 113.

Tarski’s High School Identities

Stanley Burris and Simon Lee

1. THE HIGH SCHOOL IDENTITIES. There are eleven basic identities of the positive integers which one learns in high school, namely

$$\text{HSI} \left\{ \begin{array}{l} \overline{\text{HSI}} \left\{ \begin{array}{l} (1) \quad x + y \approx y + x \\ (2) \quad x + (y + z) \approx (x + y) + z \\ (3) \quad x \cdot 1 \approx x \\ (4) \quad x \cdot y \approx y \cdot x \\ (5) \quad x \cdot (y \cdot z) \approx (x \cdot y) \cdot z \\ (6) \quad x \cdot (y + z) \approx (x \cdot y) + (x \cdot z) \end{array} \right. \\ \hline \begin{array}{l} (7) \quad 1^x \approx 1 \\ (8) \quad x^1 \approx x \\ (9) \quad x^{y+z} \approx x^y \cdot x^z \\ (10) \quad (x \cdot y)^z \approx x^z \cdot y^z \\ (11) \quad (x^y)^z \approx x^{y \cdot z} \end{array} \end{array} \right.$$

An algebra $\langle A, +, \times, \uparrow, 1 \rangle$ which satisfies HSI is called an **HSI-algebra**; and an algebra $\langle A, +, \times, 1 \rangle$ which satisfies $\overline{\text{HSI}}$ is called an **$\overline{\text{HSI}}$ -algebra**. \mathbf{N} is the HSI-algebra $\langle N, +, \times, \uparrow, 1 \rangle$, where N is the set of positive integers, and $+$, \times , \uparrow are the familiar operations of addition, multiplication and exponentiation; and $\overline{\mathbf{N}}$ is the familiar $\overline{\text{HSI}}$ -algebra $\langle N, +, \times, 1 \rangle$.

One is so accustomed to associating the identities HSI with the natural numbers that it is perhaps surprising to discover that (i) there are lots of finite HSI-algebras, and (ii) there are identities true of \mathbf{N} which cannot be derived from HSI. The question of whether or not HSI provides a basis for all the identities of \mathbf{N} has a history going back at least to the 1960s, and is called **Tarski’s High School Problem**. The negative solution to this problem was first given in 1980 by Wilkie in [10] where he shows that the identity

$$\begin{aligned} & \left((1+x)^y + (1+x+x^2)^y \right)^x \cdot \left((1+x^3)^x + (1+x^2+x^4)^x \right)^y \\ & \approx \left((1+x)^x + (1+x+x^2)^x \right)^y \cdot \left((1+x^3)^y + (1+x^2+x^4)^y \right)^x \end{aligned}$$

which we call $W(x, y)$, holds in \mathbf{N} , but cannot be derived from HSI.

To see that Wilkie’s identity actually holds consider the fact that $p(x) = 1 - x + x^2$ maps N into N , and that $1 + x^3 = p(x)(1 + x)$ and $1 + x^2 + x^4 = p(x)(1 + x + x^2)$. By pulling a factor of $p(x)^{xy}$ out of each of the two sides of Wilkie’s

identity we see that the two sides are obviously equal. (Note that we are able to prove that Wilkie's identity holds with the help of an auxiliary operation $p(x)$ which is not available in the High School Identities.)

In §6 we will give an elementary proof of the other half of Wilkie's result, namely we show that one cannot derive $W(x, y)$ from HSI.

2. THE IDENTITIES OF \bar{N} ; AND OF N

Theorem 2.1. *The identities true of \bar{N} are axiomatized by \overline{HSI} , and they are decidable.*

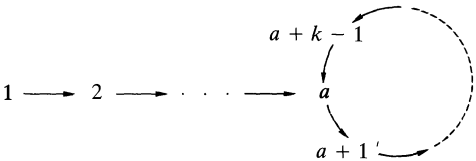
This theorem is based on the use of \overline{HSI} to obtain nice *normal forms* (namely polynomials). To test if \bar{N} satisfies an equation $s \approx t$ just express s , respectively t , as polynomials s' , respectively t' . Then $s \approx t$ holds iff $s' = t'$, i.e., we have the same polynomial.

Unfortunately there is no simple, workable notion of a normal form with respect to HSI; and HSI does not axiomatize the identities of N . Indeed Gurevič [5] showed there is no finite set of identities that axiomatizes the identities of N . (Henson and Rubel [6] used Nevanlinna theory to show that an interesting portion of these identities could be derived from HSI.) Macintyre [7] proved that the identities of N are decidable; an alternate proof was given by Gurevič [4]. Both proofs depend on sophisticated results in analysis, studying the behavior of derivatives. (The first uses results of G. H. Hardy, the second results of A. G. Hovanski.) Gurevič also gives a very simple proof in [4] (using quotients of finitely generated free HSI-algebras) that an equation is not a consequence of HSI iff it fails on a finite HSI-algebra. This shows one can also decide if an identity follows from HSI.

3. FINITE QUOTIENTS OF N ; AND PRIME NUMBERS. Quotients of N give an obvious way to construct finite models of HSI. Recall that the finite quotients of $\bar{Z} = \langle Z, +, \times, 1 \rangle$ are given by \bar{Z}/\equiv_k , where \equiv_k is *equivalent modulo k* . A variation on this leads to the finite quotients of N . For $a, k \in N$ define $\equiv_{a,k}$ on N by

$$m \equiv_{a,k} n \Leftrightarrow m = n; \text{ or } a \leq m, n \text{ and } m \equiv_k n.$$

The finite quotients of N are the $N/\equiv_{a,k}$, where $a, k \in N$ are such that $\equiv_{a,k}$ preserves exponentiation, i.e., we want $m \equiv_{a,k} n \Rightarrow x^m \equiv_{a,k} x^n$ for all $x \in N$. Let $N_{a,k} = N/\equiv_{a,k}$ (when this is defined). One can visualize $N_{a,k}$ as a k -loop with a tail of length $a - 1$:



Just as one shows that \bar{Z}/\equiv_k satisfies the identities of \bar{Z} (identities are preserved by quotients), one can show that the $N_{a,k}$ satisfy all the identities of N . In particular the $N_{a,k}$'s will satisfy the Wilkie identity.

Thanks to a communication from Peter Hoffman we have the following.

Theorem 3.1. *The finite quotients of \mathbf{N} are the $\mathbf{N}_{a,k}$ where $a, k \in \mathbf{N}$ satisfy (for all primes p):*

$$p^e | k \Rightarrow e \leq a$$

$$p | k \Rightarrow (p - 1) | k.$$

Let $\mathbf{N}_{a,k}$ be a quotient of \mathbf{N} with $k > 1$, and let $k = p_1^{e_1} \cdots p_r^{e_r}$ with $p_1 < \cdots < p_r$. As $p_1 - 1 | k$ and all the prime factors of k are larger than $p_1 - 1$ it follows that $p_1 - 1 = 1$; thus $p_1 = 2$. And if $r > 1$ then $p_2 - 1 | k$ leads to $p_2 - 1 | p_1^{e_1}$, so p_2 is of the form $2^m + 1$, and thus it is a Fermat prime. The next corollary gives a complete list of the five “circle” integer HSI-algebras, i.e., those with $a = 1$, and hence no “tail.”

Corollary 3.2. [D. Higgs]. $\mathbf{N}_{1,k}$ is a quotient of \mathbf{N} iff $k \in \{1, 2, 6, 42, 1806\}$.

Jeff Shallit pointed out that the related sequence 2, 3, 7, 43, 1807 occurs in a number of places in the literature, e.g., as solutions to Sylvester’s recurrence equations (see Davison & Shallit [3], Sylvester [8], [9]).

Given $a \in \mathbf{N}$ define the sequence of primes $\Sigma_a = (p_1, p_2, \dots)$ by

- $p_1 = 2$;
- given p_1, \dots, p_i , let p_{i+1} be the smallest prime p which is greater than p_i and such that $(p - 1) | (p_1 \cdots p_i)^a$, assuming such a p exists. If no such p exists then Σ_a terminates with p_i .

Proposition 3.3. *Given a positive integer a , there are infinitely many $\mathbf{N}_{a,k}$ iff the sequence of primes Σ_a is infinite.*

By Higgs’s result $\Sigma_1 = (2, 3, 7, 43)$, a finite sequence. However it is not known if Σ_a is finite for all (any) $a > 1$. We found that about 20% of the primes below 1,000,000 are in $\Sigma_2 = (2, 3, 5, 7, 11, 13, 19, 23, \dots, 999667, 999727, \dots)$ —so even if Σ_2 is finite, a computer enumeration does not look feasible.

4. SOME CONSTRUCTIONS OF HSI-ALGEBRAS. We have seen one easy way to construct finite HSI-algebras, by taking quotients of \mathbf{N} . Next we give five ways to construct finite HSI-algebras which have apparently nothing to do with \mathbf{N} . In four of the cases below we let exponentiation be the *first projection* function π (defined by $\pi(a, b) = a$).

- Let $\mathbf{H} = \langle H, \vee, \wedge, \rightarrow, 0, 1 \rangle$ be a Heyting algebra.
Then $\mathbf{H}^\star = \langle H, \vee, \wedge, \leftarrow, 1 \rangle$ is an HSI-algebra, where $a \leftarrow b$ is defined to be $b \rightarrow a$.
- Let $\mathbf{D} = \langle D, \vee, \wedge, 1 \rangle$ be a distributive lattice with 1.
Then $\langle D, \vee, \wedge, \pi, 1 \rangle$ is an HSI-algebra.
- Let $\mathbf{S} = \langle S, \wedge, 1 \rangle$ be a semilattice with 1.
Then $\langle S, \wedge, \wedge, \pi, 1 \rangle$ is an HSI-algebra.
- Let $\mathbf{S} = \langle S, \wedge, 0, 1 \rangle$ be a semilattice with 0, 1.
Then $\langle S, f, \wedge, \pi, 1 \rangle$ is an HSI-algebra, where f is the binary constant map whose value is always 0.
- Let $\mathbf{R} = \langle R, +, \times, 0, 1 \rangle$ be a Boolean ring.
Then $\langle R, +, \times, \pi, 1 \rangle$ is an HSI-algebra.

Each of these five constructions yields a single 2-element HSI-algebra (no two of these five are isomorphic), and these are, up to isomorphism, the only 2-element HSI-algebras. At present we do not know if all 2-element HSI-algebras satisfy all identities of \mathbf{N} . However they do satisfy Wilkie's identity (as does every HSI-algebra of size ≤ 6).

5. COMPUTER ENUMERATION. Using a computer we determined that, up to isomorphism, there are 44 3-element HSI-algebras. One can also enumerate all 4- and 5-element HSI-algebras in a reasonable amount of time (programming in C and working on a Sun Workstation). But for six elements it begins to look doubtful—we attempted to do this, and after letting the program run for several days we estimated that it could take six months of CPU time to finish. For seven elements computer enumeration seems hopeless.

6. THE WILKIE IDENTITY. $W(x, y)$ is the simplest identity known which holds on \mathbf{N} , but cannot be derived from HSI. Wilkie gave a syntactic proof that $W(x, y)$ cannot be derived from HSI; later, in 1985, Gurevič [4] published a proof by constructing a 59-element algebra satisfying HSI but not $W(x, y)$. In Gurevič [5], p. 30, we have the remark:

C. W. Henson once asked if there are countermodels to Tarski's question (whether all valid identities in signature $(+, \cdot, \uparrow)$ were derivable) of a very small size, say, 5. Currently I don't know; my own record was 33 elements and I heard a rumour that someone had pushed the record further to 28 elements.

We define a *Gurevič-algebra* (or *G-algebra*) to be a model of HSI that does not satisfy Wilkie's identity $W(x, y)$. Five is indeed too small a size for a G-algebra, as is 6. The smallest G-algebra has at least 7 elements, a result established in [1]. A 28-element G-algebra was found by Burris in 1988, and then a 17-element example in 1990. Shortly after this Simon Lee found a 16-element example; and just recently he found the smallest G-algebra known, with 15 elements:

+	1	2	3	4	a	b	c	d	e	f	g	h	i	j	k
1	2	3	4	4	2	3	d	3	4	4	4	4	4	4	4
2	3	4	4	4	3	4	3	4	4	4	4	4	4	4	4
3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
a	2	3	4	4	b	3	b	3	4	4	4	4	4	4	4
b	3	4	4	4	3	4	3	4	4	4	4	4	4	4	4
c	d	3	4	4	b	3	b	3	4	4	4	4	4	4	4
d	3	4	4	4	3	4	3	4	4	4	4	4	4	4	4
e	4	4	4	4	4	4	4	4	4	g	4	4	4	4	4
f	4	4	4	4	4	4	4	4	g	4	4	i	4	4	4
g	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
h	4	4	4	4	4	4	4	4	4	i	4	4	4	4	4
i	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
j	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
k	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

\times	1	2	3	4	a	b	c	d	e	f	g	h	i	j	k
1	1	2	3	4	a	b	c	d	e	f	g	h	i	j	k
2	2	4	4	4	b	4	b	4	4	4	4	4	4	4	4
3	3	4	4	4	3	4	3	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
a	a	b	3	4	c	b	c	b	4	f	4	4	4	4	4
b	b	4	4	4	b	4	b	4	4	4	4	4	4	4	4
c	c	b	3	4	c	b	c	b	4	f	4	4	4	4	4
d	d	4	4	4	b	4	b	4	4	4	4	4	4	4	4
e	e	4	4	4	4	4	4	4	4	4	4	j	4	4	4
f	f	4	4	4	f	4	f	4	4	4	4	4	4	4	4
g	g	4	4	4	4	4	4	4	4	4	4	4	4	4	4
h	h	4	4	4	4	4	4	4	j	4	4	4	4	4	4
i	i	4	4	4	4	4	4	4	4	4	4	4	4	4	4
j	j	4	4	4	4	4	4	4	4	4	4	4	4	4	4
k	k	4	4	4	4	4	4	4	4	4	4	4	4	4	4

\uparrow	1	2	3	4	a	b	c	d	e	f	g	h	i	j	k
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	4	4	4	e	4	4	4	4	4	4	4	4	4	4
3	3	4	4	4	f	4	4	4	4	4	4	4	4	4	f
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
a	a	c	c	c	c	c	c	c	c	c	c	c	c	c	c
b	b	4	4	4	4	4	4	4	4	4	4	4	4	4	4
c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
d	d	4	4	4	4	4	4	4	4	4	4	4	4	4	h
e	e	4	4	4	4	4	4	4	4	4	4	4	4	4	4
f	f	4	4	4	4	4	4	4	4	4	4	4	4	4	4
g	g	4	4	4	4	4	4	4	4	4	4	4	4	4	h
h	h	4	4	4	4	4	4	4	4	4	4	4	4	4	4
i	i	4	4	4	e	4	4	4	4	4	4	4	4	4	4
j	j	4	4	4	4	4	4	4	4	4	4	4	4	4	4
k	k	4	4	4	4	4	4	4	4	4	4	4	4	4	4

This HSI-algebra is such that $W(x, y)$ fails only for the pair $(x, y) = (a, k)$. The verification that this is an HSI-algebra, and that $W(a, k)$ does not hold, constitutes a proof that Wilkie's identity does not follow from the High School Identities. The eight-element $\overline{\text{HSI}}$ -algebra $1, 2, 3, 4, a, b, c, d$ (boxed in above) was found with the help of a computer. The rest was done by hand.

So the smallest HSI-algebra which rejects $W(x, y)$ has between 7 and 15 elements. These are the best bounds we know.

REFERENCES

1. S. Burris and S. Lee, Small models of the High School Identities, *Intermat. J. Alg. and Computation*, **2** (1992), 139–178.
2. S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra*, Grad. Texts in Math, **78**, Springer-Verlag, 1981.
3. J. L. Davison and J. O. Shallit, Continued fractions for some alternating series, *Monatshefte Math.*, **111** (1991), 119–126.

4. R. Gurevič, Equational theory of positive numbers with exponentiation, *Proc. Amer. Math. Soc.*, **94** (1985), 135–141.
5. R. Gurevič, Equational theories of positive numbers with exponentiation is not finitely axiomatizable, *Annals of Pure and Applied Logic* **49** (1990), 1–30.
6. C. W. Henson and L. A. Rubel, Some applications of Nevanlinna theory to mathematical logic: identities of exponential functions, *Trans. Amer. Math. Soc.* **282** (1984), 1–32.
7. A. Macintyre, The laws of exponentiation, *Springer Lecture Notes in Math.*, **890** (1981), 185–197.
8. J. J. Sylvester, On a point in the theory of vulgar fractions, *Amer. J. Math.* **3** (1880), 332–335.
9. J. J. Sylvester, Postscript to a note on a point in vulgar fractions, *Amer. J. Math.* **3** (1880), 388–389.
10. A. J. Wilkie, On exponentiation—a solution to Tarski’s high school algebra problem, preprint, Oxford University, 1980.

Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1

Department of Mathematics
Columbia University
New York, NY 10027

Last winter a questionnaire was sent to 50 leading universities in America asking for information concerning persons who already held the doctorate or probably would secure it during 1934 and who were seeking positions for 1934–35. Nearly all of the universities replied, and 120 persons were named who were seeking positions. There were 60 other men and women who received the doctorate in mathematics during 1934. Many of these 180 people held positions, some of which might be considered permanent but others were certainly temporary.

•
•
•

The situation might be roughly summarized by stating that there were about 40 or 50 Doctors of Philosophy in mathematics who had not, on October first, found employment reasonably satisfactory to them.

—*American Mathematical Monthly*
 42, (1935) p. 143.

Aperiodic Chaotic Orbits

Steven N. MacEachern and L. Mark Berliner

1. INTRODUCTION. Two common formalizations of the popular notion of unpredictability in the presence of chaos are sensitive dependence on initial conditions and expansiveness. Sensitive dependence on initial conditions suggests that orbits of chaotic systems corresponding to even very nearby initial conditions *may* separate as time grows. Expansiveness occurs when all initial conditions yield orbits which must, for some time value, be separated by some minimum distance $\delta > 0$. This note sharpens the distinction between these concepts.

The definition of a discrete time, dynamical system acting on a compact set, J , of the real line involves a function, say f , which maps J into itself. The function is applied iteratively. Specifically, the value of the system at time n is $x_{n+1} = f(x_n) = f^{n+1}(x)$, where $x \in J$ and $f^n(\cdot)$ denotes the n -fold composition of f . An orbit of the dynamical system is denoted by $\{x_n\}$. For ease of exposition, the presentation here is specialized to a familiar example for $J = [0, 1]$, namely, the tent map defined by the function

$$\begin{aligned} f(x) &= 2x & 0 \leq x < .5 \\ &2 - 2x & .5 \leq x \leq 1. \end{aligned} \tag{1.1}$$

Generalizations of much of the analysis to other continuous unimodal maps from $[0, 1]$ onto $[0, 1]$ for which $f(0) = f(1) = 0$ are possible. In particular, the results apply to the logistic map where $f(x) = 4x(1 - x)$.

We follow Devaney [2] for the basic definitions. A dynamical system is said to exhibit sensitive dependence on initial conditions if there exists $\delta > 0$ such that, for any $x \in J$ and any neighborhood N of x , there exists $y \in N$ and $n \geq 0$ such that $|f^n(x) - f^n(y)| > \delta$. A system is said to be expansive if there exists $\delta > 0$ such that, for any $x, y \in J$ with $x \neq y$, there exists n such that $|f^n(x) - f^n(y)| > \delta$. The distinction between these two definitions is that the first claims the existence of nearby initial conditions that have orbits that separate by at least δ while the second claims that all nearby initial conditions have orbits that eventually separate by at least δ .

It is well-known that the cases of the tent map and the logistic map defined above both display sensitivity to initial conditions. However, simple consideration of initial conditions x and $1 - x$, where x is arbitrarily close to $.5$, demonstrates that neither map is expansive. In general, there exist points which eventually lead to identical orbits. The motivation of this paper is the following question: If we eliminate all such initial conditions (i.e., points which lead to orbits that are eventually identical), must the orbits corresponding to any two initial conditions eventually diverge? The answer is no.

This tent map provides an excellent illustration of the issue. On one hand, points that are on the same side of $.5$ are pulled apart with each iteration of the

function. The distance between two such points doubles with each iteration, leading to an exponential rate of separation. On the other hand, points that are on opposite sides of .5 may be brought closer together or even mapped to the same image by the folding (or “2–1” nature) of the function. Thus there is a fundamental tension between these two properties of the function. For some pairs of points, stretching dominates so that the orbits will separate by some $\delta > 0$ infinitely often. For other pairs the folding behavior is more powerful, leading to orbits which coincide after some time. The key point of this paper is that a third possibility exists as well. Namely, the case where the folding dominates, not through the coincidence of orbits after some time, but through successive folds each of which squeezes the orbits very close together. Although this case may exist, it is in some sense counterintuitive, as noted by Collet and Eckmann [1], p. 136. For pertinent discussion, see Li and Yorke [4] and Jackson [3], pp. 178–179.

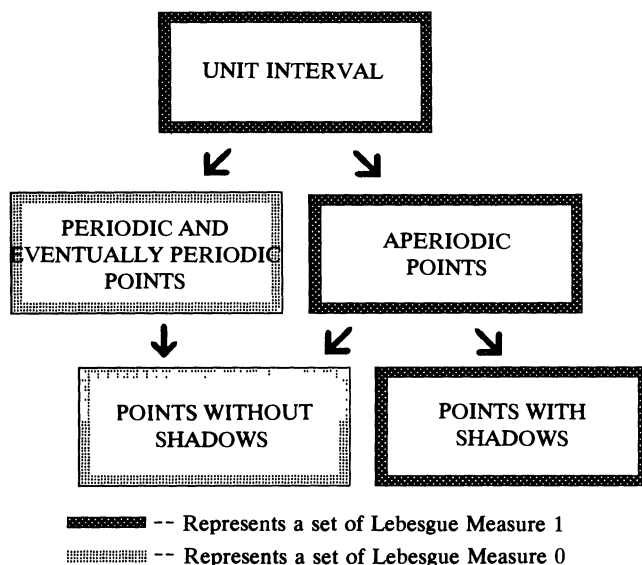


Figure 1. The study of periodicity of orbits first partitions the unit interval into two sets: The Aperiodic Points and the Periodic or Eventually Periodic Points. The Aperiodic Points have measure 1. A *shadow* of a given aperiodic point yields an orbit which never coincides with the shadowed orbit, but converges to that orbit. Theorem 1 shows that the aperiodic points which have shadows form a set of measure 1. It also shows that only aperiodic points are shadowable.

To clarify the results presented, consider the classification of points in the unit interval based on the behavior of their orbits. (See Figure 1.) The unit interval, a set of Lebesgue measure 1, is first split into two sets; namely the sets of points leading to periodic, or eventually periodic, orbits and the points leading to aperiodic orbits. It is easy to see that the former set is countable. Hence, the aperiodic points form a set of Lebesgue measure 1. For an aperiodic point, x , define a *shadow* of x to be any point y leading to an orbit for which $|x_n - y_n| > 0$, for all n , but $\lim_{n \rightarrow \infty} |x_n - y_n| = 0$. (Note that we are using a very strong definition of a shadow. We do not consider the notion of shadowing a chaotic orbit by a numerically computed version.) Theorem 1 shows that the set of aperiodic points which have shadows is also of Lebesgue measure 1. While this result may

seem to suggest that sensitivity to initial conditions is not as evident as we might expect, Theorem 2 vindicates our intuition by showing that, for a given aperiodic, shadowable point, the set of its shadows has Lebesgue measure 0.

The results of this paper rely on the notion of symbolic dynamics. Specifically, we construct symbolic dynamic representations of orbits which are aperiodic and do not coincide, yet which converge to each other. The richness of the maps considered here insures that actual orbits of the system corresponding to these symbolic dynamic sequences exist.

2. MAIN RESULTS. The following definitions and results are used in the development of our results. First, we turn to the topic of symbolic dynamics. See Devaney [2] and Collet and Eckmann [1] for motivation of this tool. For an orbit $\{x_n\}$, with initial condition $x_0 = x$, define another sequence $\mathbf{t}_x = \{t_0, t_1, t_2, \dots\}$ where

$$\begin{aligned} t_i &= 0 & \text{if } x_i < .5 \\ &1 & \text{if } x_i \geq .5. \end{aligned} \tag{2.1}$$

The sequence \mathbf{t}_x will be referred to as the symbolic dynamic sequence of the orbit $\{x_n\}$. Let $\mathbf{T} = \{\mathbf{t} : \text{there exists } x \in J \text{ with } \mathbf{t} = \mathbf{t}_x\}$. An additional definition used below is the *shift map* σ applied to symbolic dynamic sequences where $\sigma(t_0, t_1, t_2, \dots) = t_1, t_2, \dots$. In our case \mathbf{T} contains all infinite sequences of 0's and 1's such that $\sigma^j(\mathbf{t})$ is not the sequence with a 0 in position 1, a 1 in position 2, and 0's thereafter for any j (see Devaney [2], Section 1.18).

In the context of the tent map, several features of the relationship between $\{x_n\}$ and \mathbf{t}_x should be noted. First, the function (2.1) creates a 1-1 map from the orbits $\{x_n\}$ onto \mathbf{T} . Second, an orbit $\{x_n\}$ is periodic if and only if there exists $j > 0$ such that $\sigma^j(\mathbf{t}_x) = \mathbf{t}_x$. This property also implies that an orbit which is eventually periodic (i.e., that settles into a periodic orbit after some initial transitory behavior has washed out) satisfies $\sigma^{n+j}(\mathbf{t}_x) = \sigma^n(\mathbf{t}_x)$ for some $n, j > 0$. Third, two sequences coincide after some time point (i.e. if $x_n = y_n$ for some n) if and only if $\sigma^n(\mathbf{t}_x) = \sigma^n(\mathbf{t}_y)$ for some n . Taken together, these properties state that if two sequences have symbolic dynamic representations that differ infinitely often and that are not eventually repeating, then they are aperiodic and will not coincide.

By examining consecutive entries of the symbolic sequences \mathbf{t}_x and \mathbf{t}_y , we may determine about how close the corresponding orbits are to each other. These distances are best examined by means of the m -segment.

Definition. The m -segment of a symbolic dynamic sequence \mathbf{t} is the finite sequence $\{t_0, \dots, t_m\}$ and is denoted $\mathbf{t}(m)$.

An ordering is imposed on the set of potential m -segments by the map. Since the map is continuous and 2-1 onto J , all of the potential m -segments will be represented. These 2^{m+1} m -segments correspond to a partition of J into 2^{m+1} intervals. If two m -segments correspond to adjacent intervals we say the m -segments are adjacent. A picture for the case of $m = 3$ appears below. Reading down a column gives the first four entries of the symbolic dynamic for points in each 3-segment. The width of each interval in $[0, 1]$ corresponding to an m -segment is 2^{-m-1} . Thus, if the two orbits with initial conditions x and y yield symbolic dynamics with $\mathbf{t}_x(m) = \mathbf{t}_y(m)$, then $|x - y| \leq 2^{-m-1}$; if $\mathbf{t}_x(m)$ and $\mathbf{t}_y(m)$ are adjacent, then $|x - y| \leq 2^{-m}$.

Theorem 1. *For the tent map (1.1), the set of points which have a shadow has Lebesgue measure 1.*

Proof: Let A be the set of points which have a shadow. Let B be the set of points which have arbitrarily long, but not infinitely long, strings of 0's in their symbolic sequences. Let C be the set of points for which $\{f^n(x)\}$ has $1/2$ as a limit point. We will show that $A = B = C$.

The set B has measure 1. Since two iterations of the tent map send a point near $1/2$ to an image near 0, we easily see that $B = C$.

The following argument shows that A is contained in C . Let y be a shadow of x . Then $|f^n(x) - f^n(y)| > 0$ for all n , and so there exists a subsequence n_k for which $f^{n_k}(x)$ and $f^{n_k}(y)$ are on opposite sides of $1/2$. But $\lim_{n_k \rightarrow \infty} |f^{n_k}(x) - f^{n_k}(y)| = 0$, and so $\lim_{n_k \rightarrow \infty} f^{n_k}(x) = 1/2$.

Next, we show that B is contained in A . Let $L_1 = 0$, $L_2 = 01$, $L_3 = 010$, $L_4 = 0100$, and $L_5 = 01000$. Let $R_1 = 1$, $R_2 = 11$, $R_3 = 110$, $R_4 = 1100$, and $R_5 = 11000$. More generally, let $L_m = 010 \dots 0$ and $R_m = 110 \dots 0$, where the two sequences are of length m and have $m - 2$ trailing 0's. The width of the interval of points which have for their first m symbols L_m (i.e., the $(m - 1)$ -segment just to the left of $.5$) is 2^{-m} . The width of the interval of points which have for their first m symbols R_m (i.e., the $(m - 1)$ -segment just to the right of $.5$) is 2^{-m} . (See Figure 2.)

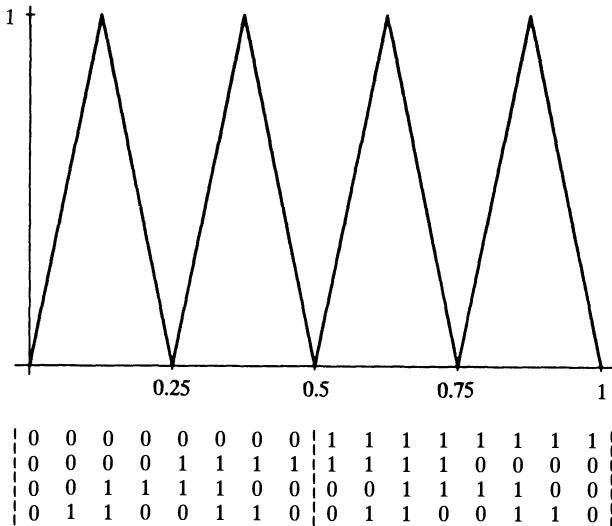


Figure 2. Three iterates of the tent map, as plotted, partition the unit interval into 16 subintervals. The points in these subintervals all have the indicated symbols, (t_0, t_1, t_2, t_3) , in the first four components of their symbolic sequences.

Lemma. *If x has first m symbols $e_1, e_2, \dots, e_k, L_{m-k}$ and y has first m symbols $e_1, e_2, \dots, e_k, R_{m-k}$ then $|x - y| \leq 2^{1-m}$.*

This is because $f^k(x)$ and $f^k(y)$ have the adjacent $(m - k - 1)$ -segments determined by L_{m-k} and R_{m-k} , respectively. Then $|f^k(x) - f^k(y)| \leq 2^{1-(m-k)}$.

$f^{k-1}(x)$ and $f^{k-1}(y)$ are inverse images of $f^k(x)$ and $f^k(y)$, and they are on the same side of $1/2$ so $|f^{k-1}(x) - f^{k-1}(y)| = |f^k(x) - f^k(y)|/2 \leq 2^{1-(m-(k-1))}$. Repeating this argument gives $|x - y| \leq 2^{1-m}$, establishing the Lemma.

Now suppose x is in B . Then the symbol sequence for x is of the form $\text{string}_1\{L_3 \text{ or } R_3\}\text{string}_2\{L_4 \text{ or } R_4\}\text{string}_3\{L_5 \text{ or } R_5\}\dots$, where string_k is a sequence of 0's and 1's of length ≥ 0 . (Note that this form is not uniquely determined, because string_k can have any length.) Choose y to have the same symbol sequence as x , but switch the L_i 's with R_i 's, and vice-versa. Then y is a shadow of x .

Note that the definition of the set B implies that only aperiodic points are shadowable. The following theorem shows that the Lebesgue measure of the set of initial conditions that correspond to convergent sequences is 0.

Theorem 2. *Let x be any point in the unit interval. For the tent map (1.1), the set of all shadows of x has Lebesgue measure 0.*

Proof: Let S_0 denote the set of all shadows of x . We first show that S_0 is measurable. Let $S_{m,k} = \{y: \sigma^k(t_y)(m) \text{ be adjacent to or identical to } \sigma^k(t_x)(m)\}$. $S_{m,k}$ is the union of 2^k intervals, and is therefore measurable. Since

$$S_0 = \bigcap_{m=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} S_{m,k}$$

S_0 is Lebesgue measurable. Let $\mu(S_0)$ denote the Lebesgue measure of S_0 .

Let $S_1 = f(S_0)$ denote the set of all shadows of $f(x)$. For any $y \in S_1$, with $y \neq 1$, there are exactly two values of z (namely, $z = y/2$ or $1 - y/2$) for which $\sigma(t_z) = t_y$. (The point $y = 1$ is mapped to 0 and, therefore, cannot shadow any point.) Since the map (1.1) is symmetric, linear, and doubles lengths of subintervals of $[0, .5]$ and of $[.5, 1]$, $\mu(S_0 \cap [0, .5]) = \mu(S_0 \cap [.5, 1]) = \mu(S_1)/2$. This implies $\mu(S_1) = \mu(S_0)$. Repeating this argument for the set of shadows of $f^k(x)$, $S_k = f^k(S_0)$, we obtain $\mu(S_k) = \mu(S_0)$ for each $k > 0$.

To see that S_k and S_j are disjoint, consider the case $j > k$. As noted in the proof of Theorem 1, there exists a subsequence n_i such that $\lim_{n_i \rightarrow \infty} f^{n_i+k}(x) = .5$. This implies that $\lim_{n_i \rightarrow \infty} f^{n_i+j}(x) = 0$; that is, a point cannot shadow both $f^j(x)$ and $f^k(x)$.

ACKNOWLEDGMENTS. We gratefully acknowledge the contributions of a Referee. In particular, Theorem 1 is a strengthened version, due to the Referee, of our original.

REFERENCES

1. P. Collet and J.-P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Birkhauser, Berlin, 1980.
2. R. L. Devaney, *Introduction to Chaotic Dynamical Systems* (2/e), Addison-Wesley, Menlo Park, 1989.
3. E. A. Jackson, *Perspectives of Nonlinear Dynamics*, Cambridge University Press, New York, 1989.
4. T. Y. Li and J. A. Yorke, Period three implies chaos, *Amer. Math. Monthly* (82) (1975) 983–992.

*Department of Statistics
Ohio State University
Columbus, OH 43210*

Bricklaying and the Hermite Normal Form

William J. Gilbert

We describe a geometric interpretation of the Hermite normal form of an integer matrix in terms of tilings by bricks.

A non-singular integer $n \times n$ matrix A generates an integer lattice, L , in n -space. The lattice points are the integer linear combinations of the columns of A . Figure 1 shows such a lattice in the plane.

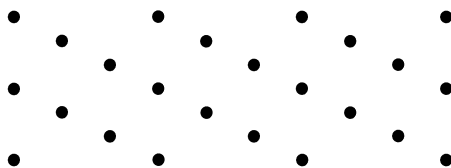


Figure 1. A lattice generated by the columns of $\begin{bmatrix} 2 & 2 \\ -1 & 2 \end{bmatrix}$.

How can we tile n -space by rectangular bricks parallel to the coordinate axes so that the translation group of the tiling is the same as that of the lattice L ? Figures 2 and 3 show two such tilings with the same translation group as the lattice in Figure 1. The lower left corner of each brick lies on a lattice point.

Some Basic Definitions

An *integer lattice* in n -dimensional space is the set of all integer linear combinations of n linearly independent vectors with integer entries.

The *translation group of the lattice* consists of all the translations of n -dimensional space that take lattice points to lattice points. This group is isomorphic to the set of lattice points under vector addition.

A *fundamental brick* will be a rectangular brick with one vertex at the origin and sides along each positive coordinate axis. A *brickwork* will consist of a tiling of n -space by n -dimensional bricks that are translations of a fundamental brick. The set of these translations form the *translation group of the brickwork*.

Each lattice has many different bases. We can change the basis of the lattice L by means of unimodular column operations on the matrix A . The *elementary unimodular column operations* are essentially “integer” column operations and

consist of

- (i) adding an integer multiple of one column to another;
- (ii) exchanging two columns;
- (iii) multiplying one column by -1 .

Performing unimodular column operations on a matrix is equivalent to post-multiplying by a *unimodular* matrix U ; i.e. an integer matrix U whose determinant is ± 1 . The inverse of U is also an integer matrix.

Every integer matrix can be reduced, by unimodular column operations, to a unique standard form called the Hermite normal form [1, Ch. I.7], [2, Part II]. A non-singular integer matrix is said to be in *Hermite normal form* if it is a lower triangular, non-negative matrix in which each row has a unique maximum entry that is located on the main diagonal. (Beware that some authors call the reduced row echelon form over a field the Hermite form. Even those authors and computer packages that define the Hermite normal form over the integers, or a Euclidean ring, do so in different ways. These are all equivalent, but some use row operations instead of column operations and some use upper triangular matrices instead of lower triangular ones.)

For example, the lattice in Figure 1 is generated by the matrix $A = \begin{bmatrix} 2 & 2 \\ -1 & 2 \end{bmatrix}$ and its Hermite normal form is $H = \begin{bmatrix} 2 & 0 \\ 2 & 3 \end{bmatrix}$. This matrix H can be considered as corresponding to the brickwork in Figure 2. The diagonal elements determine the size of the bricks. The last column determines the edge along the last coordinate, while the first column determines the offset of an adjacent brick. In general, the columns of H determine the lattice points on the boundary of the fundamental brick.

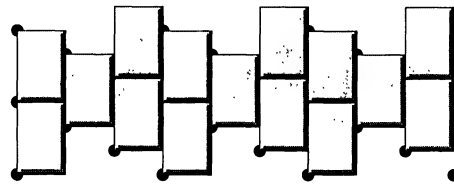


Figure 2. Bricks of size 2×3 corresponding to the matrix $\begin{bmatrix} 2 & 0 \\ 2 & 3 \end{bmatrix}$.

The Hermite normal form is obtained by unimodular reduction and so the magnitude of the determinant is unchanged. This magnitude is the area of a generating parallelogram of the lattice (or volume of a generating parallelepiped) and also the area (or volume) of a brick.

What is the connection between the brickwork in Figure 3 and the Hermite normal form? The algorithm for reducing a matrix to Hermite normal form initially consists of column reducing the matrix so that the first row is in the required form with all the off-diagonal entries zero. This first leading entry is then the greatest common divisor of all the entries in the first row. The algorithm continues inductively by leaving the first r columns alone and then column reducing the others so that the last $(n - r - 1)$ entries in the $(r + 1)$ st row are zero. Then multiples of the $(r + 1)$ st column are added to the first r columns so that the $(r + 1)$ st row is in the required form.

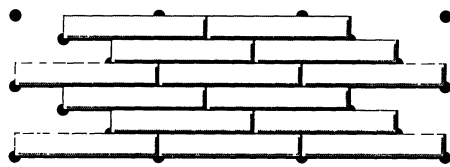


Figure 3. Bricks of size 6×1 corresponding to the matrix $\begin{bmatrix} 6 & 4 \\ 0 & 1 \end{bmatrix}$.

If we start with the matrix A and modify this algorithm so that the second row has all off-diagonal entries zero, then we obtain the matrix $\begin{bmatrix} 6 & 4 \\ 0 & 1 \end{bmatrix}$, that corresponds to the brickwork in Figure 3.

For an $n \times n$ matrix A , there are $n!$ ways of modifying the Hermite normal form algorithm, corresponding to the $n!$ different orderings of the rows. Let π be a permutation of the n rows of A . If we inductively leave the columns $\pi^{-1}(1), \dots, \pi^{-1}(r)$ alone and column reduce the matrix so that the other off-diagonal entries of the $\pi^{-1}(r+1)$ st row are zero, we obtain a modified Hermite normal form, H_π . This corresponds to a brickwork with the same translation group as that derived from A . H_π is a non-negative matrix whose (i, j) th entry is zero whenever $\pi(i) < \pi(j)$ and is such that each row has a unique maximum entry located on the main diagonal. If P is the permutation matrix corresponding to π , then this modified Hermite normal form is

$$H_\pi = P^{-1} \text{hnf}(PA)P$$

where $\text{hnf}(PA) = PAU$, the Hermite normal form of PA , and U is the unimodular matrix corresponding to the column operations. Notice that $H_\pi = AUP$ where UP is a unimodular matrix. Since $PH_\pi P^{-1} = \text{hnf}(PA)$ is lower triangular, it follows that the (i, j) th entry of H_π is zero whenever $\pi(i) < \pi(j)$.

Figure 4 shows a typical tiling of 3-space by bricks corresponding to the Hermite normal form of a 3×3 matrix. There are six ways of modifying this Hermite normal form. The following is an example in which the six ways give rise to brickworks that all use different sized bricks.

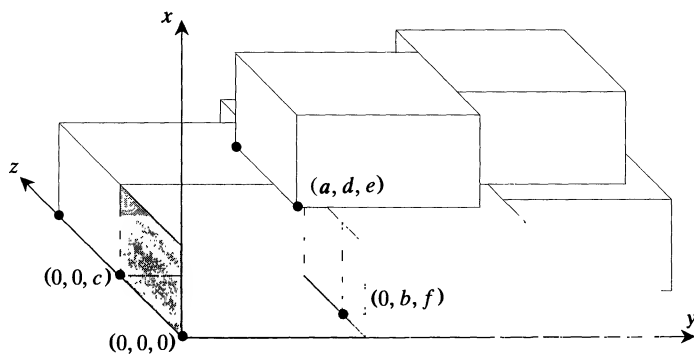


Figure 4. Bricks of size $a \times b \times c$ corresponding to the matrix $\begin{bmatrix} a & 0 & 0 \\ d & b & 0 \\ e & f & c \end{bmatrix}$.

Consider the matrix

$$A = \begin{bmatrix} 5 & 2 & 1 \\ -4 & 2 & 4 \\ 0 & -3 & 6 \end{bmatrix}.$$

If we perform the modified algorithm described above using the six permutations of the rows, we obtain the six matrices

$$H_{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 6 & 0 \\ 6 & 15 & 30 \end{bmatrix}, \quad H_{(123)} = \begin{bmatrix} 5 & 0 & 3 \\ 8 & 12 & 6 \\ 0 & 0 & 3 \end{bmatrix}, \quad H_{(132)} = \begin{bmatrix} 15 & 8 & 6 \\ 0 & 2 & 0 \\ 0 & 3 & 6 \end{bmatrix},$$

$$H_{(12)} = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 2 & 0 \\ 18 & 27 & 30 \end{bmatrix}, \quad H_{(13)} = \begin{bmatrix} 15 & 10 & 8 \\ 0 & 4 & 2 \\ 0 & 0 & 3 \end{bmatrix}, \quad H_{(23)} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 12 & 6 \\ 6 & 0 & 15 \end{bmatrix}.$$

These matrices all induce the same lattice in 3-space and they correspond to six brickworks that all use different sized bricks, namely $1 \times 6 \times 30$, $5 \times 12 \times 3$, $15 \times 2 \times 6$, $3 \times 2 \times 30$, $15 \times 4 \times 3$, and $1 \times 12 \times 15$.

All the brickworks having the same translation group as that derived from an integer matrix A arise from the modified Hermite normal forms applied to A . This essentially follows from the Minkowski-Hajós problem on tilings by n -cubes and induction on the dimension n . In 1907, H. Minkowski considered a problem in number theory that was equivalent to the following geometric conjecture. These problems are described in [3].

Minkowski's Conjecture. *If a lattice of unit n -cubes tiles n -space, then some pair of cubes share a complete $(n - 1)$ -dimensional face.*

In 1942, G. Hajós proved the conjecture for all n by changing it into an equivalent conjecture concerning finite abelian groups.

By shrinking each axis separately, any brickwork tiling n -space using rectangular bricks is affinely equivalent to a tiling by unit n -dimensional cubes. Hence some pair of bricks must share a complete $(n - 1)$ -dimensional face. This $(n - 1)$ -dimensional face is perpendicular to one of the axes, say the j th axis. If e_j is the unit vector along this axis, then one of the edges of the bricks is along the vector $a_j e_j$, for some positive integer a_j . The lattice point $a_j e_j$ is a vertex of the fundamental brick and every brick lies in a complete row of bricks parallel to the j th axis, that is obtained by translating that brick by integer multiples of $a_j e_j$.

Now project the tiling onto the $(n - 1)$ -dimensional subspace orthogonal to the j th axis. A similar argument using the Minkowski-Hajós problem in $(n - 1)$ -space will show that another lattice point on the fundamental brick is of the form $b_k e_k + b_j e_j$, where $k \neq j$, b_k and b_j are integers, $b_k > 0$, and $0 \leq b_j < a_j$.

Use induction on the dimension to complete the basis for the lattice, consisting of lattice points on the boundary of the fundamental brick. This basis will form the columns of one of the modified Hermite normal forms obtained from A .

REFERENCES

1. G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, New York, 1988.
2. A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, Chichester, 1986.
3. S. K. Stein, Algebraic Tiling, *Amer. Math. Monthly*, 81 (1974) 445–462.

Pure Mathematics Department
University of Waterloo
Waterloo, Ontario
CANADA N2L 3G1

A Characterization of Inner Product Spaces

Neil Falkner

Let \mathcal{A} be a vector space over the field \mathbb{R} of real numbers. (We consider the complex case later.) Suppose $a \mapsto \|a\|$ is a function from \mathcal{A} to \mathbb{R} . Let us say that $a \mapsto \|a\|$ is a *Euclidean norm* when it satisfies the following two conditions:

- E1.** For each $a \in \mathcal{A}$ and each $t \in \mathbb{R}$, $\|ta\| = |t| \|a\| \geq 0$ and if $\|a\| = 0$, then $a = 0$;
E2. For each $a, b \in \mathcal{A}$ and each $t > 1$, if $\|a\| = \|b\|$, then $\|a - tb\| = \|b - ta\|$.

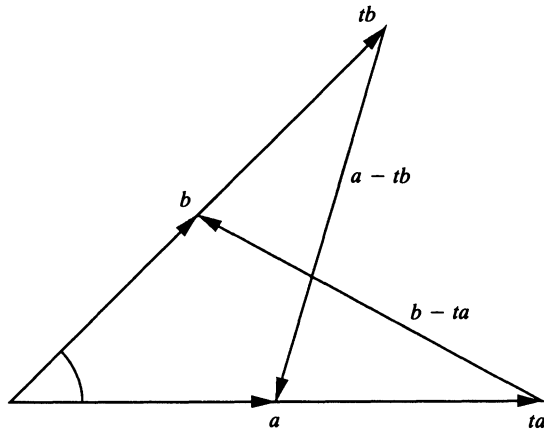


Figure 1.

The object of this note is to show that if $a \mapsto \|a\|$ is a Euclidean norm, then it is the norm arising from an inner product on \mathcal{A} . (The converse is obvious.) It follows that a Euclidean norm does satisfy the triangle inequality, although this was not assumed in the definition. (We remark that in Euclid, the triangle inequality is deduced, not assumed; see [7, Book I, Proposition 20].)

Note that E2 may be regarded as asserting that the side-angle-side rule of congruence for triangles holds in the very special case where the two corresponding angles are not merely congruent but actually coincide. See Figure 1. The restriction $t > 1$ is included in E2 only to make this geometric interpretation as clear as possible. It follows trivially from E1 and E2 that if $\|a\| = \|b\|$, then $\|ua + vb\| = \|va + ub\|$ for each $u, v \in \mathbb{R}$. Ficken [5] showed that if the latter condition holds in a normed linear space, then the norm arises from an inner product. Our result improves on his in that we do not assume the triangle inequality. In addition, the proof we present is shorter than his. We wish to thank the referees for suggesting improvements in the presentation.

We assume henceforth that E1 and E2 both hold. In Euclidean geometry, a right angle is defined to be one which is congruent to its supplementary angle. With this in mind, let us write $c \perp d$ to mean $\|c - d\| = \|c - (-d)\|$.

Observe that for each $c, d \in A$, and each $r, s \in \mathbb{R}$ if $c \perp d$, then $rc \perp sd$. In addition, it is worth mentioning that, under E1, this property of \perp is actually equivalent to E2. These assertions follow from the fact that the system of equations

$$\begin{aligned}a &= c - d \\b &= c + d \\ua + vb &= rc + sd \\va + ub &= rc - sd\end{aligned}$$

can be solved for the quantities on either side in terms of the quantities on the other side and furthermore $\|a\| = \|b\|$ if and only if $c \perp d$.

Now consider any two-dimensional subspace B of A . It follows from the previous paragraph that B has an “orthonormal” basis e, f . Observe that $xe + yf \perp -ye + xf$ for each $x, y \in \mathbb{R}$. For let $c = xe + yf$ and $d = -ye + xf$. Then

$$\begin{aligned}\|c - d\| &= \|(x + y)e - (x - y)f\| \\&= \|(x + y)e + (x - y)f\| \quad (\text{since } e \perp f) \\&= \|(x - y)e + (x + y)f\| \quad (\text{since } \|e\| = \|f\|) \\&= \|c + d\|.\end{aligned}$$

Hence $c \perp d$ as claimed. Next, we prove the “Pythagorean theorem”: $\|xe + yf\|^2 = x^2 + y^2$ for each $x, y \in \mathbb{R}$. We may assume that $y \neq 0$. Let $c = xe + yf$, $r = (x^2 + y^2)^{1/2}$, and $d = re$. We wish to show that $\|c\| = \|d\|$. Now $c = a + b$ and $d = a - b$ where $2a = c + d = (x + r)e + yf$ and $2b = c - d = (x - r)e + yf$. Note that $2b = w[-ye + (x + r)f]$ where $w = (r - x)/y$. Thus $a \perp b$ so $\|c\| = \|d\|$ as desired. It follows from the “Pythagorean theorem” that on B , $\|\cdot\|$ is the norm arising from the inner product $(\cdot | \cdot)_B$ given by $(x_1e + y_1f | x_2e + y_2f)_B = x_1x_2 + y_1y_2$. From this it follows that $(\cdot | \cdot)_B$ agrees on $B \times B$ with the function $(\cdot | \cdot)$ defined on $A \times A$ by

$$(g|h) = \frac{1}{4}(\|g + h\|^2 - \|g - h\|^2). \quad (1)$$

Note that $(g|h) = 0$ if and only if $g \perp h$.

We have now seen that $(\cdot | \cdot)$ defined by (1) is an inner product on each two-dimensional subspace of A . In particular,

$$(a|b + c) = (a|b) + (a|c) \quad (2)$$

whenever a, b , and c are linearly dependent. The parallelogram-law $\|h + e\|^2 + \|h - e\|^2 = 2(\|h\|^2 + \|e\|^2)$ follows. It remains only to show that (2) holds even when a, b , and c are linearly independent. But this follows from the parallelogram-law by a well-known calculation due to Jordan and von Neumann [8] which is a generalization of the main idea of the proof of Proposition 4 in Book XI of Euclid [7] and which may best be understood with the aid of a picture. Let $d = (b + c)/2$ and $e = (b - c)/2$ so that $d + e = b$ and $d - e = c$.

In the parallelogram-law, instead of thinking of a parallelogram, think of a triangle with median h and adjacent sides $h + e$ and $h - e$, so $\|e\|$ is half the

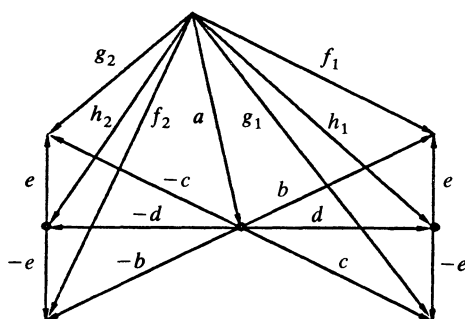


Figure 2.

length of the side bisected by h . Referring to Figure 2, we have

$$\begin{aligned}
 4(a|b) + 4(a|c) &= (\|f_1\|^2 - \|f_2\|^2) + (\|g_1\|^2 - \|g_2\|^2) \\
 &= (\|f_1\|^2 + \|g_1\|^2) - (\|f_2\|^2 + \|g_2\|^2) \\
 &= 2(\|h_1\|^2 + \|e\|^2) - 2(\|h_2\|^2 + \|e\|^2) \\
 &= 2(\|h_1\|^2 - \|h_2\|^2) = 8(a|d) = 4(a|b + c).
 \end{aligned}$$

(The cited proposition in Euclid states that if three distinct lines meet at a point and if one of them is perpendicular to each of the other two, then it is perpendicular to the plane containing these two.)

Now suppose A is actually a vector space over the complex field \mathbb{C} . We continue to assume that E1 and E2 hold. Let $[a|b] = (ia|b)$. Suppose that $\|ia\| = \|a\|$ for each $a \in A$. (This is equivalent to each of the following two conditions: (i) $[a|a] = 0$ for each $a \in A$; (ii) $[b|a] = -[a|b]$ for each $a, b \in A$.) Then it is well-known and easy to check that the formula $\langle a|b \rangle = (a|b) + i[a|b]$ defines a complex inner product on A (\mathbb{C} -linear in its second argument) which gives rise to the norm $\|\cdot\|$. In particular, we have $\|za\| = |z| \|a\|$. It is worth noting that two vectors a and b are orthogonal in the complex sense (i.e., $\langle a|b \rangle = 0$) if and only if the planes $\mathbb{C}a$ and $\mathbb{C}b$ are orthogonal in the real sense (i.e., $(za|wb) = 0$ for each $z, w \in \mathbb{C}$).

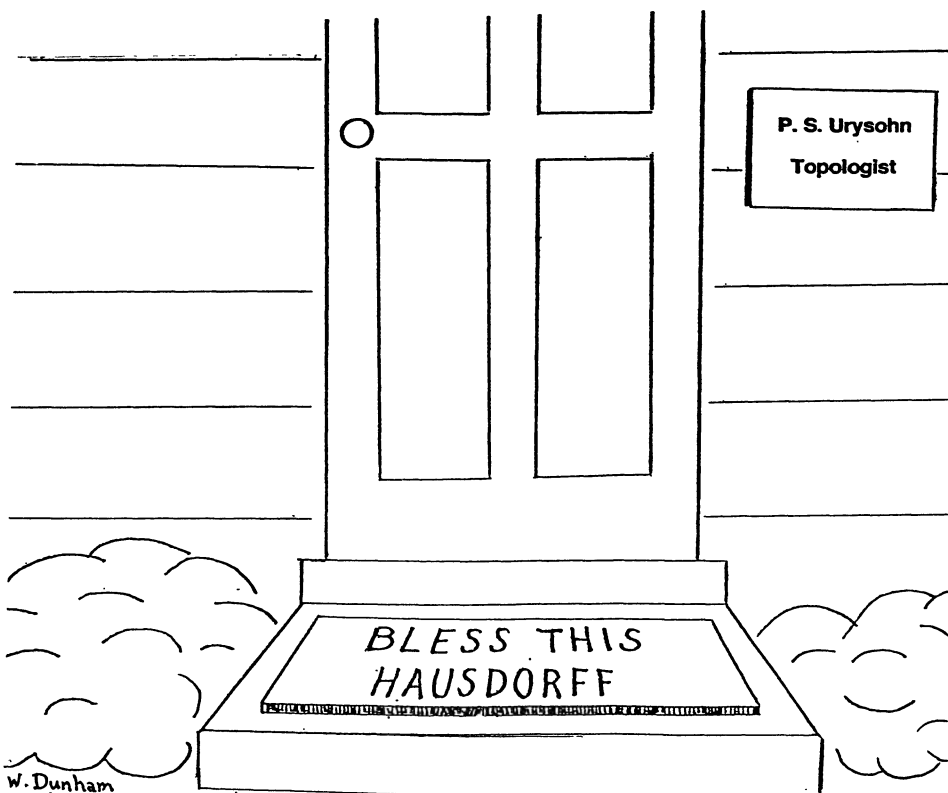
This note may be viewed as treating the metric aspect of the derivation of analytic geometry from synthetic geometry. Another aspect of this derivation is the construction of the vector space itself from more primitive geometric material. This construction is elegantly described in [2] for the two-dimensional case. Discussions of the three-dimensional case may be found in [3] and [6] while the case of three or more dimensions is treated in [9]. Reference [4] is an interesting supplement to [9].

REFERENCES

1. Dan Amir, *Characterizations of Inner Product Spaces*, Birkhäuser, 1986.
2. E. Artin, *Geometric Algebra*, Wiley, 1957.
3. R. P. Burn, *Deductive Transformation Geometry*, Cambridge University Press, 1975.
4. Francis Bukenhout, Une caractérisation des espaces affins basée sur la notion de droite, *Mathematische Zeitschrift*, 111 (1969) 367–371.

5. Frederick A. Ficken, Note on the existence of scalar products in normed linear spaces, *Annals of Mathematics*, (2) 45 (1944) 362–366.
6. Günter Ewald, *Geometry: An Introduction*, Wadsworth, 1971.
7. Sir Thomas L. Heath, *The Thirteen Books of Euclid's Elements*, 3 volumes, second edition, Cambridge University Press, 1926; Dover reprint, 1956.
8. P. Jordan and J. von Neumann, On inner products in linear metric spaces, *Annals of Mathematics*, (2) 36 (1935) 719–723.
9. Hanfried Lenz, Ein Kurzer Weg zur analytischen Geometrie, *Mathematisch-Physikalische Semesterberichte zur Pflege des Zusammenhangs von Schule und Universität*, 6 (1958–59) 57–67.

Department of Mathematics
The Ohio State University
231 West 18th Avenue
Columbus, OH 43210



Submitted by William Dunham, Muhlenberg College.

Polar Area Is the Average of Strip Areas

Gilbert Strang

1. INTRODUCTION. Calculus finds the area of a plane region by cutting it into small pieces. This note is about the three shapes that we use most often—vertical strips, horizontal strips, and polar triangles. Those lead to $\int y dx$ and $\int x dy$ and $\frac{1}{2} \int r^2 d\theta$, for the areas marked 1, 2, 3 in the figures. It was a total surprise to learn that these three areas are directly related, even for *finite* strips and triangles.

Figure 1 shows the regions when $y(x)$ is a linear function. In that case the three shapes have straight boundaries—we have two trapezoids and a triangle. The unexpected relation is this:

The polar area is the average of the areas of the two strips.

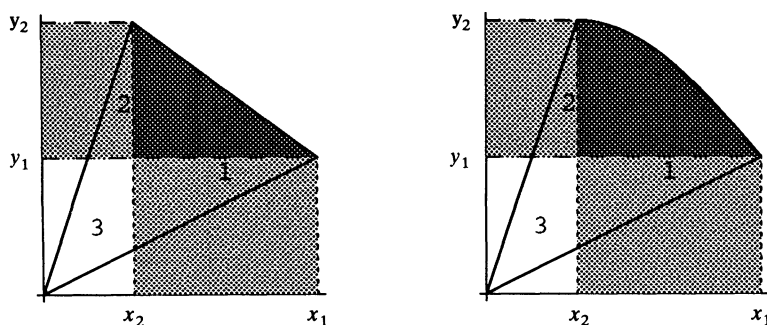


Figure 1–2. Area of triangle 3 = $\frac{1}{2}$ (area of strip 1 + area of strip 2).

The relation is still true when the function is nonlinear and its graph curves downward (Figure 2). In that case there is a natural urge to cut up the polar triangle and match it to the strips; this proof is the best. The straight-sided case has an algebraic proof, the curved case has a geometric proof, and the general case (when $y(x)$ or $x(y)$ or $r(\theta)$ may not be single valued) has a calculus proof.

So far I have no reference for the observation that $\text{area } 3 = \frac{1}{2}(\text{area } 1 + \text{area } 2)$. Certainly it cannot be new. The discussion of the three proofs leads to related questions about areas, and a clearer understanding of $\frac{1}{2} \oint (x dy - y dx)$.

Algebraic proof. The areas of the trapezoids are $\frac{1}{2}(x_1 - x_2)(y_1 + y_2)$ and $\frac{1}{2}(x_1 + x_2)(y_2 - y_1)$. The terms involving $x_1 y_1$ and $x_2 y_2$ cancel in the sum. The average of the two areas is $\frac{1}{2}(x_1 y_2 - x_2 y_1)$. We recognize this formula for the area of the triangle—it is half of the 2 by 2 determinant that gives the area of a parallelogram. This is half the length of the cross product $(x_1, x_2, 0) \times (y_1, y_2, 0)$.

This algebraic proof could be applied to thin strips with bases Δx and Δy . By combining straight-sided regions and taking the limit, curves are allowed. A geometric approach is more elementary and direct.

Geometric proof: Draw the extra line to (x_2, y_1) in Figure 3. Triangle A and rectangle R have the same height y_1 and the same base. (The base is actually their common top.) Therefore the area of A is half of the rectangular area.

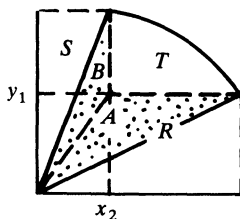


Figure 3. Polar area $|A| + |B| + |T| = \frac{1}{2}(|R| + |T| + |S| + |T|)$.

For triangle B and rectangle S at the upper left, the conclusion is the same: one area is half the other. This leaves triangle T at the top right, which is shared by all three regions—whether it is curved or straight. Then A , B , and T combine to give half the area of R , S , and $2T$. Therefore $\text{area } 3 = \frac{1}{2}(\text{area } 1 + \text{area } 2)$.

2. LINE INTEGRALS. Areas also appear later in calculus, usually as a test case for Green's Theorem. The integral of $1 \, dx \, dy$ over a plane region (which yields its area) is equal to a line integral around the boundary. That integral can be $-\oint y \, dx$ or $+\oint x \, dy$ or $\frac{1}{2}\oint(x \, dy - y \, dx)$. The third integral looks more “balanced,” but it is rarely given a meaning of its own. The main interest is in the minus sign for $\oint y \, dx$, and the application of Green's Theorem. We want to recognize our three areas.

Calculus proof: Integrate $-y \, dx$ around the vertical strip (region 1, straight or curved). The lower three sides have $dx = 0$ or $y = 0$. The minus sign is needed because the integral goes from (x_1, y_1) counterclockwise to (x_2, y_2) . The end result is the familiar calculation $\int_a^b y \, dx$. The line integral also explains [1, page 566] why areas below the x axis count as negative.

Similarly either x or dy is zero along three sides of the horizontal strip (region 2). The line integral around the boundary reduces to the usual formula $\int x \, dy$. Now integrate $\frac{1}{2}(x \, dy - y \, dx)$ along the triangle (region 3). The sides that enter and leave the origin have $dy/dx = y/x$; on those edges the integrand is zero. In all cases the only contribution to the line integrals (which equal the three areas) is along the common boundary at the top. Since $\frac{1}{2}(x \, dy - y \, dx)$ is the average of $x \, dy$ and $-y \, dx$ on that boundary, the triangle area is the average of the strip areas.

The whole picture is suggesting that $\frac{1}{2}\oint(x \, dy - y \, dx)$ is the polar area $\frac{1}{2}\int r^2 \, d\theta$. **That is true.** Substituting $x = r \cos \theta$ and $y = r \sin \theta$, the integrals agree. Note that dy includes $\sin \theta \, dr$ as well as $r \cos \theta \, d\theta$; the dr terms cancel to leave $\frac{1}{2}\int r^2(\cos^2 \theta + \sin^2 \theta) \, d\theta$.

It is tempting to associate this with a famous application of the line integral, when the ellipse $x = a \cos t$, $y = b \sin t$ has area $\frac{1}{2}\int_0^{2\pi} ab(\cos^2 t + \sin^2 t) \, dt = \pi ab$. But t is not θ ! This resolves a paradox that makes a good exercise in calculus: for

the ellipse $x = 4 \cos \theta$ and $y = 3 \sin \theta$, the area is $\pi ab = 12\pi$ but

$$\frac{1}{2} \int r^2 d\theta = \frac{1}{2} \int (x^2 + y^2) d\theta = \frac{1}{2} \int_0^{2\pi} (16 \cos^2 \theta + 9 \sin^2 \theta) d\theta = 12\frac{1}{2}\pi.$$

Important Note. Suppose $y(x)$ is not monotonically decreasing. Figure 4 shows two examples, one linear and the other curved. The linear case is really extreme: *The polar area is zero.* The two trapezoidal areas are equal. (Geometric proof: Each trapezoid is a big triangle minus a small triangle. The big triangles are equal halves of a rectangle and so are the small triangles.) Polar area is now half the *difference* in rectangular areas.

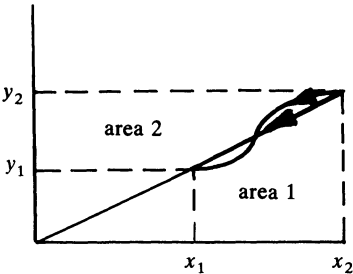


Figure 4. Positive slope leads to difference of areas.

To keep signs straight, follow the calculus proof—the line integrals. Counter-clockwise around area 1 is now clockwise around area 2. The curved path in Figure 4 goes partly clockwise and partly counterclockwise around the polar area. No wonder students have trouble with polar area; authors do too. Can a reader rescue the geometric proof (for difference of area), when $y(x)$ and $r(\theta)$ are increasing?

3. THE AREA OF A TRIANGLE. Earlier we met the formula $\frac{1}{2}(x_1 y_2 - x_2 y_1)$ for the area of triangle 3. This is familiar and correct (as the geometrical proof showed). Slightly less familiar is the 3 by 3 determinant that gives the area when the third corner is (x_3, y_3) instead of $(0, 0)$. This determinant has a mysterious *column of ones*. The direct proof is again by algebra using the rules for determinants:

$$\text{area} = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{aligned} & +\frac{1}{2}(x_1 y_2 - x_2 y_1) \\ & +\frac{1}{2}(x_2 y_3 - x_3 y_2) \\ & +\frac{1}{2}(x_3 y_1 - x_1 y_3). \end{aligned}$$

This is a sum of three “smaller” triangles. The first one connects $(0, 0)$ to (x_1, y_1) and (x_2, y_2) . The other triangles go from $(0, 0)$ to the other two edges in Figure 5. The three areas match the three terms from the determinant, and add to the area of the whole triangle T .

The sum still gives the correct area in Figure 6 when $(0, 0)$ is outside the triangle and two areas have negative signs. We could also divide T into trapezoids!

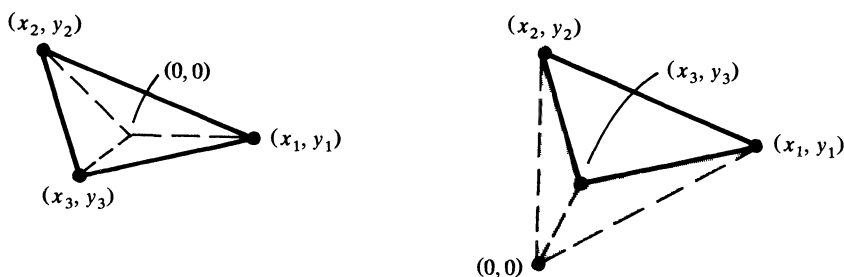


Figure 5-6. Triangle = sum (or difference) of three triangles from $(0, 0)$.

A more geometrical proof can motivate that column of 1's. To get into three dimensions, lift the triangle T up to the plane $z = 1$. The corners are now $(x_1, y_1, 1)$ and $(x_2, y_2, 1)$ and $(x_3, y_3, 1)$. The triangle is the top of an upside-down pyramid in Figure 7, coming out from $(0, 0, 0)$. Therefore the pyramid volume is $\frac{1}{3}$ of the area of T (the pyramid height is 1). This pyramid is also $\frac{1}{6}$ of a box, so its volume is $\frac{1}{6}$ of our determinant. Here is the link between volume (geometry) and the determinant (algebra). Then $\frac{1}{3}$ of the area of T equals $\frac{1}{6}$ of the determinant, and multiplying by 3 gives the area formula with the factor $\frac{1}{2}$.

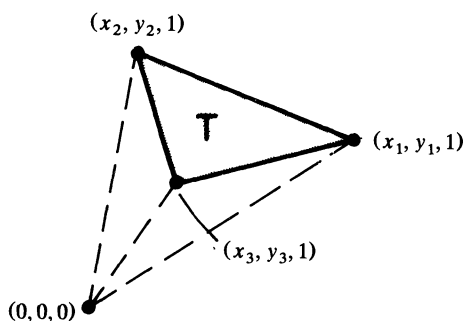


Figure 7. Pyramid volume gives area of T .

4. THE AREA OF A POLYGON. Barry Cipra told me about meeting two carpet-layers in a Minnesota coffee shop. They wanted the area of a polygonal carpet, knowing its corners. Inspired by caffeine he proposed

$$\text{area of } P = \frac{1}{2}[(x_1 y_2 - x_2 y_1) + (x_2 y_3 - x_3 y_2) + \cdots + (x_n y_1 - x_1 y_n)]. \quad (1)$$

This formula is correct (and known but not well known). The simplest proof is to connect each edge to the origin. Then P is a union of triangles, as in Figure 8. Again triangles with reverse orientation contribute negative area to the sum. Another proof is to integrate $\frac{1}{2}(x dy - y dx)$ along each edge in turn, and use Green's Theorem. This differential could be seen directly as half the length of a cross product, between the radial vector (x, y) and the step (dx, dy) along the outer edge. That is again the area of a thin polar triangle.

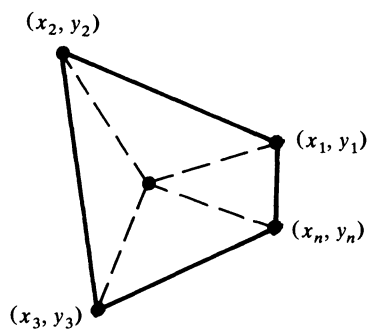


Figure 8. Polygon into triangles.

Bart Braden has given a beautiful discussion [3] that develops the connection to polar areas and Green's Theorem. He refers to equation (1) as the *surveyor's area formula*. Surveyors also write the area of P as $\frac{1}{2}\sum x_i(y_{i+1} - y_{i-1})$.

My own experience was with a lawyer, who wanted to know the area of a property. "If I count steps walking around the boundary, what is the area?" He was very disappointed in mathematicians when I needed to know the shape. Now I would tell him the carpet-layer formula (without mentioning line integrals). Still he was aiming in the right direction, to solve a two-dimensional problem by working with its one-dimensional boundary.

In three dimensions or higher, the lawyer and the carpet-layers would be left behind—and we are deeper into the serious calculus of differential forms.

REFERENCES

1. Gilbert Strang, *Calculus*, Wellesley-Cambridge Press, Box 812060, Wellesley MA (1991).
2. Barry Cipra, private communication.
3. Bart Braden, The surveyor's area formula, *College Math. Journal*, 17 (1986) 326–337.

Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139
gs@math.mit.edu

The search for truth is more precious than its possession.

—Einstein

Counting Critical Points of Real Polynomials in Two Variables

Alan Durfee, Nathan Kronenfeld, Heidi Munson, Jeff Roy,
and Ina Westby

The graph of a real function of two variables can be thought of as a landscape with hills, valleys, lakes and passes. If the lakes are drained and the landscape is an average one, the only level spots will be the tops of the hills, the bottom of the lakes, and the passes. Suppose now that the function is a polynomial. The problem is to count the numbers of peaks, pits and passes, and relate these numbers to the degree of the polynomial.

This topic was investigated in 1989 at the ten-week Summer Research Institute in Mathematics at Mount Holyoke College, where the first version of this paper was written. Four undergraduates between their junior and senior years collaborated on the project. They were directed by Alan Durfee of the mathematics faculty at Mount Holyoke. The Institute was supported by the Research Experiences for Undergraduates program of the National Science Foundation, and the New England Consortium for Undergraduate Science Education. The topic was suggested by V. I. Arnold of Moscow State University especially for our research group. This apparently simple problem turned out to be quite complicated, and led us in many interesting directions. We especially thank Arnold for giving us such an inspiring topic.

In more mathematical terms, the problem is the following: Given a real polynomial of two variables with only nondegenerate critical points, relate the numbers of local maxima, minima and saddles of the polynomial to its degree. There are two aspects of this problem, examples and theory. First, the examples: We searched through polynomials of low degree, looking for patterns. We looked at perturbations of a polynomial with a degenerate critical point. We looked in calculus books, but these contained only simple polynomials suitable for hand computation. We did, however, find an inspiring example in a journal article, a polynomial with just one critical point, a local maximum, but not an absolute maximum (see the end of Section 6).

How does one find information about a polynomial starting with its equation? We used computer graphics packages (*Mathematica* and *MacFunction*) to make contour plots and three-dimensional graphics of our polynomials. This required skill. An inappropriate scale for the range of a 3D-picture of a polynomial may show a flat space where in fact there are two summits and a saddle. Contour plots have the same problem. However, graphics did help us in understanding our task. A second method was to use computer algebra (again, *Mathematica*) to find the common zeros of the two partial derivatives of the polynomial and hence the critical points. We mechanized the whole process by programming the computer to

take a polynomial and return a list of maxima, minima and saddles. However, this method was obviously limited by the speed of the computer.

Next, the theory: The relevant results are well-known and appealingly simple. A critical point occurs at a common zero of the two partials of the polynomial. Since the zero locus of each partial is an algebraic curve, the number of critical points is bounded by Bezout's Theorem, which states that the number of intersections of two curves in the real plane is at most the product of their degrees. Another bound comes from topology: The index of the gradient vector field of the polynomial at a nondegenerate critical point is plus or minus one, depending on whether the critical point is a local extremum or a saddle. Also, the absolute value of the index of this vector field on a large circle containing all the critical points is less than the degree of the polynomial. Combining these two bounds gives inequalities on the number of saddles and the number of extrema.

One interesting class of polynomials are those which factor into real linear factors. The zero set of such a polynomial is an arrangement of lines, which we assume are in general position. The polynomial has a saddle point where the lines intersect, and each bounded region of the arrangement contains exactly one local extremum. We discuss an upper bound for the number of local maxima of such polynomials, and some examples [Ch], [FP] of arrangements where this bound is attained.

There is an obvious connection between our problem and Hilbert's Sixteenth Problem on the arrangements of ovals of real algebraic curves: The zero set of a polynomial is a collection of ovals, and each oval must have a local extremum in its interior. Ragsdale [Ra] gives examples of polynomials of even degree with a large number of positive ovals (ovals on whose immediate interior the polynomial is positive) and thus for which the number of maxima is roughly three times the number of minima.

We added to the theory by showing that if the gradient vector field of a polynomial has index greater than one around a large circle, then the polynomial must have multiple critical points at infinity. (A critical point at infinity is a point on the line at infinity in projective space where the completions of the zero loci of the two partials intersect.) The proof involves looking carefully at the geometry of these gradient vector fields. We obtain as a corollary a better upper bound on the number of extrema. These results are due to Jeff Roy.

There is a large gap between the above theory and the examples. The theory restricts the number of maximums, minimums and saddles that a polynomial can have, yet there are few examples to show that these restrictions are optimal. Even though we were able to contribute results to the theory and add more polynomials to the collection of examples, the gap remains as wide as ever.

1. PRELIMINARY CONCEPTS. Let $f(x, y)$ be a polynomial in two variables with real coefficients. We let

$$d = \text{degree of } f$$

The polynomial has a *critical point* at (x_0, y_0) if $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$ where f_x and f_y denote the partial derivatives of the function f with respect to x and y , respectively. If

$$\det \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix} = 0$$

then (x_0, y_0) is a *degenerate critical point*; if this determinant is non-zero, then (x_0, y_0) is a *nondegenerate critical point*. The only nondegenerate critical points of a polynomial in two real variables are local maxima, minima, and saddles. All nondegenerate critical points are isolated. We let

m = number of local maxima of f

n = number of local minima of f

s = number of saddles of f .

Recall that a real *plane algebraic curve* is the zero set of a real polynomial in two variables. For background material on this topic, see for instance [BK], [Wal], [Fu], or [Gri]. A proof of the following result can be found in [BK, p. 227].

Theorem 1.1 (Bezout). *If A and B are algebraic curves in the real plane of degree a and b , respectively, and A and B have no common components, then A and B intersect in at most ab points.*

Bezout's Theorem in fact says that the two curves intersect in exactly ab points in the complex projective plane. These intersections must be counted with multiplicities. A simple algorithm for computing multiplicities can be found in [Fu].

Proposition 1.2. *A polynomial of degree d with isolated critical points has at most $(d - 1)^2$ critical points (counted with multiplicities).*

Proof: If $f(x, y)$ is a polynomial of degree d , then $f_x = 0$ and $f_y = 0$ are algebraic curves of degree at most $d - 1$. These curves have no common components since these components would consist entirely of critical points and hence not be isolated. The critical points of the polynomial occur at the intersections of these two curves, so by Bezout's theorem the total number of critical points is at most $(d - 1)^2$. \square

What distinguishes a polynomial from an arbitrary smooth function? A polynomial can have only a finite number of critical points, as demonstrated above. Another property is that all the level curves of a polynomial intersect the line at infinity at the same points, and there are at most d of them, since these intersection points are exactly the zeros of the homogeneous part of highest degree. Furthermore, a polynomial is 'unbounded' in a strong sense:

Proposition 1.3. *Let $f(x, y)$ be a polynomial of degree $d > 1$. If f is bounded on more than d lines then (after a possible rotation of coordinates) f is a polynomial in one variable.*

Proof: Suppose f is bounded on $e > d$ lines l_1, \dots, l_e . The polynomial f is constant on each of these lines, since when restricted to each line it is a polynomial of one variable which is bounded and hence constant.

If not all l_i are parallel, then f takes the same value c on all the lines. Then $l_1 l_2 \dots l_e - c$ divides f , a contradiction since $e > d$. If all the l_i are parallel, then choose x', y' such that the x' -axis is parallel to the l_i and the y' -axis is perpendicular to the l_i . Then l_i has equation $\{y' = c_i\}$. For each k , the function f restricted to $\{x' = k\}$ is determined by the e points $(k, c_1), \dots, (k, c_e)$ since $e > d$, and is in fact determined independent of k . \square

2. THE INDEX OF THE GRADIENT VECTOR FIELD. First we discuss the index of a vector field. Let γ be an oriented closed curve in the plane, and suppose that v is a nonzero smooth vector field on γ , so that each point on the curve has a vector associated to it. The index of the vector field around the curve is roughly speaking the number of counterclockwise revolutions that the vector makes while a point travels completely around the curve. More precisely:

Definition 2.1. The *index* of v around γ is the topological degree of the map $h: \gamma \rightarrow S^1$ given by $h(p) = v(p)/|v(p)|$.

The index can also be defined as the number of times that the vector field points in a fixed direction, counted with a plus or minus sign according to the way the vector field is turning at that point; more precisely, it is the number of points counted with signs in the inverse image of a regular value of h . (See [Mi, Lemma 3, p. 36 and Lemma 4, p. 37]). We will give some examples of the index below, when we discuss the gradient vector field of a polynomial. If $v(p) = 0$, the *index of v at p* is defined to be the index of v around a circle (oriented counterclockwise) containing p and no other zeros of v . The following result says that the index of a vector field around a curve is a topological quantity. In particular, it says that if the curve is deformed in a region where the vector field is nonzero, then the index remains the same. A proof can be found in [Mi] or [GP, Ch. 3, Sect. 5].

Theorem 2.2. *If γ is a closed counterclockwise-oriented curve which does not pass through a zero of the vector field v , then the index of v around γ is equal to the sum of the indices of v at all zeros interior to γ .*

How can the index of a vector field around a curve be computed? The following lemma expresses the index as the integral of the change in angle of the vector field. (Although the integral can be evaluated numerically using a computer, there may be problems if the vector field turns rapidly in a short distance, as does for example the gradient vector field of the polynomial $y^5 + y^3x^2 - y$ discussed below.)

Lemma 2.3. *The index of a vector field v around a circle γ centered at the origin is*

$$\frac{1}{2\pi i} \int_{\gamma} \frac{dv}{v}.$$

Here v is written in complex notation: if v has components v_1 and v_2 , then $v = v_1 + iv_2$.

Proof: Write $v = re^{2\pi i\theta}$, where r and θ are functions of x and y . Then $\theta = 1/(2\pi i)(\log v - \log r)$. Since r is constant on the circle γ , we have $d\theta = 1/(2\pi i)d \log v = 1/(2\pi i) dv/v$. Integrating both sides around γ proves the lemma. \square

We now look at the gradient vector field of a real polynomial $f(x, y)$.

Definition 2.4. The *index of $f(x, y)$ at an isolated critical point $p \in \mathbf{R}^2$* is the index of $\text{grad}(f)$ around a circle (oriented counterclockwise) containing p and no other critical points of f . The index of $f(x, y)$ is the index of $\text{grad}(f)$ around a circle (oriented counterclockwise) containing all the critical points of f .

We let

$$i = \text{index of } f$$

A nondegenerate maximum or minimum has index $+1$ and a nondegenerate saddle has index -1 . The index of a polynomial with nondegenerate critical points of which m are local maxima, n are local minima and s are saddles is thus

$$i = m + n - s.$$

We next prove a simple bound on the absolute value of the index of a polynomial about a circle.

Proposition 2.5. *The index i of a polynomial $f(x, y)$ of degree d about a circle C in the plane satisfies $|i| \leq d - 1$.*

Proof: The locus where the gradient vector field is horizontal (pointing either right or left) is given by the curve $D = \{(x, y) | f_y = 0\}$. The points on C where the vector field is horizontal are exactly the points where C and D intersect. If C is a component of D , the gradient on C always points in either the positive or negative x direction, and $i = 0$. If C and D have no common components, they intersect in isolated points, say j of them. By Bezout's Theorem, $j \leq 2(d - 1)$. Without loss of generality, we may assume that pointing left and right are regular values for the map h in Definition 2.1, and thus $2|i| \leq j$. This proves the proposition. \square

The index of a polynomial at a point is easy to compute using the following result, which is proved in [Ar1] using techniques from Morse theory. (The result can also be proved using the techniques of Section 6.)

Proposition 2.6. *The index of a polynomial $f(x, y)$ at a point p_0 is $1 - r$, where r is the number of real branches at p_0 of the curve $f(x, y) = f(p_0)$.*

The lower and upper bounds of Proposition 2.5 are quite different. We may easily obtain polynomials which reach the lower bound: The polynomial $f(x, y)$ defined by the real part of $(x + iy)^d$ has a degenerate critical point at the origin and it is easy to see that this polynomial has index $1 - d$. For example, if $d = 3$ we obtain the well-known 'monkey saddle'. A small perturbation by adding terms of lower degree produces a polynomial with the same index and nondegenerate critical points.

Proposition 2.6 implies that the index of a polynomial at a point can be at most one. The upper bound on the index of a polynomial around a large circle is more complicated. For example, the only critical points of the polynomial $y^5 + y^3x^2 - y$ are a minimum and a maximum, and thus the polynomial has index of $+2$. (Fig. 2.1, 2.2.) (These figures and the following ones of polynomials $f(x, y)$ are of the corresponding "scrunched" polynomial $\arctan f(\tan x, \tan y)$ for $-\pi/2 < x, y < \pi/2$, with some scaling factors as in [Sm]. The advantage of these pictures is that they show the qualitative behavior of the function $f(x, y)$ at infinity. Note, however, that "scrunching" slightly displaces the location of the critical points.)

These are also polynomials with arbitrarily high index: for instance, the polynomial

$$\prod_{k=1}^n (y(x^2 + k) - 1)$$

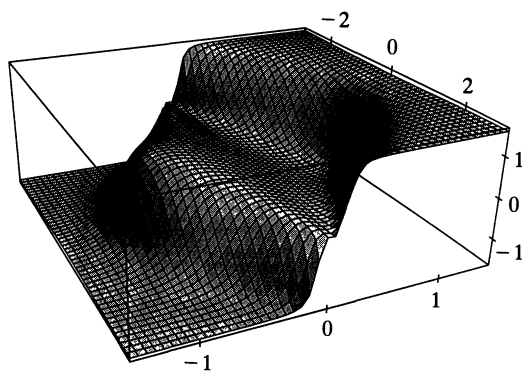


Figure 2.1. $f(x, y) = y^5 + y^3x^2 - y$

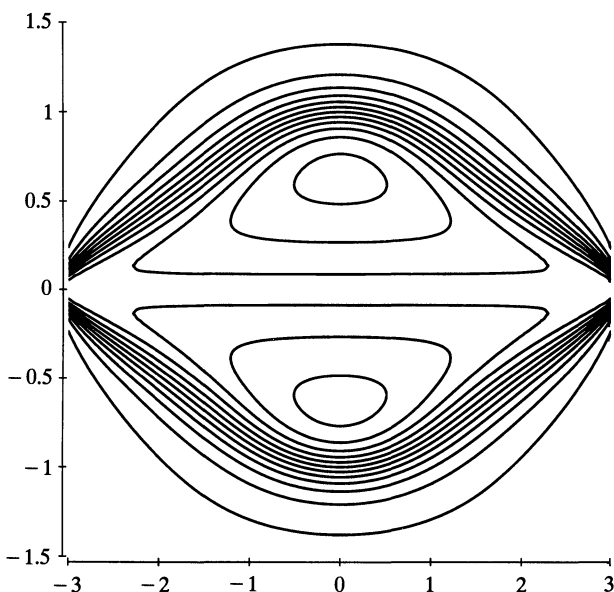


Figure 2.2. $f(x, y) = y^5 + y^3x^2 - y$

has index $n - 1$. A formula for computing the index i from information at infinity is given in [Du].

Problem 2.7. *Are there polynomials with $i = d - 1$, the upper bound of Proposition 2.5?*

The answer is probably “no.”

3. BOUNDS ON THE NUMBER OF CRITICAL POINTS. Let $f(x, y)$ be a polynomial of degree d , and suppose that f has only nondegenerate critical points. Let m be the number of local maxima, let n be the number of local minima, and let s be the number of saddle points of f as before.

First we discuss the bounds obtained by combining the results of Proposition 1.2 and Proposition 2.5. Proposition 1.2 says that

$$m + n + s \leq (d - 1)^2$$

and Proposition 2.5 says that

$$1 - d \leq m + n - s \leq d - 1.$$

From these two inequalities we get

$$s \leq \frac{1}{2}(d^2 - d) \tag{3.1}$$

and

$$m + n \leq \frac{1}{2}(d^2 - d) \tag{3.2}$$

The polynomials of Section 4 show that the bound 3.1 is sharp. We will derive a slightly better form of equation 3.2 later (Corollary 6.9). Also, 5.4 provides a conjectural bound for the number of local maxima m .

Problem 3.1. *Find more inequalities similar to the above, and find examples of polynomials whose critical points satisfy the various possibilities allowed by these inequalities.*

What about polynomials with only one type of critical point? It is relatively easy to construct polynomials with only saddles, for example by perturbing the generalized monkey saddle. Also, the two polynomials at the end of the previous section have no saddles, only local extrema. What about just local maxima? According to Mathematics Magazine [Ro], a certain calculus book asserted that it was impossible to have a function of two variables with two local maxima and no other critical points; the magazine challenged the reader to find a counterexample. A polynomial counterexample (Fig. 3.1, 3.2) was given in a later issue [Da]:

$$f(x, y) = -(x^2y - x - 1)^2 - (x^2 - 1)^2.$$

This polynomial has local maxima at the points (1, 2) and (−1, 0). It would be nice to find a polynomial f with more symmetry than this one, for instance with the property $f(x, y) = f(-x, y)$.

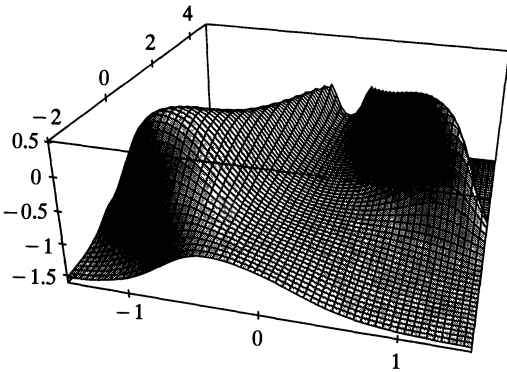


Figure 3.1. $f(x, y) = -(x^2y - x - 1)^2 - (x^2 - 1)^2$

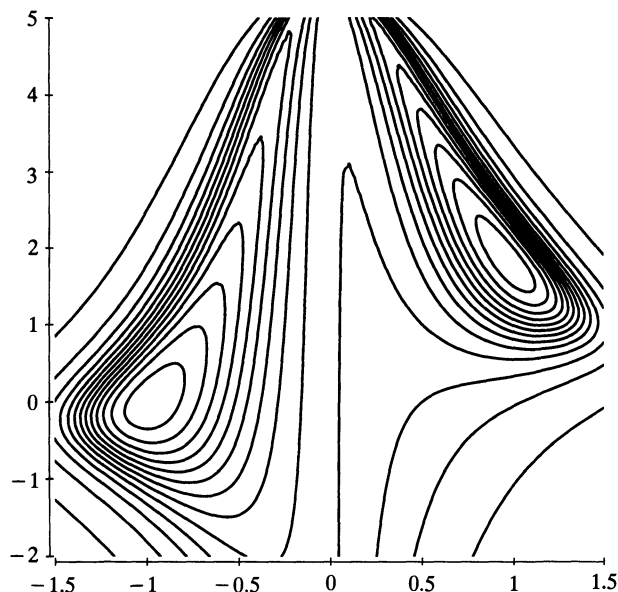


Figure 3.2. $f(x, y) = -(x^2y - x - 1)^2 - (x^2 - 1)^2$

Problem 3.2. Find polynomials whose only critical points are k nondegenerate local maxima, for $k > 2$.

There are easily constructed smooth functions whose only critical points are local maxima. For example, the function $e^{-x}(xe^{-x} + \cos y)$ has an infinite number of local maxima and no other critical points. Pictures of this function can be found in [Wag, Sect. 3.4]. Its geometry is as follows: On the line $x = 0$, the graph is the cosine. As x approaches negative infinity, the graph drops down rapidly. As x approaches positive infinity, the graph approaches the level 0; the points from the maxima gradually slope down to 0 and the points from the minima gradually slope up to 0. This type of geometry is not possible for polynomials, since this function has value 0 for $x = +\infty$ and in particular does not satisfy Proposition 1.3. In all these examples, and also those at the end of Section 6, one can also think of the saddle points as having been “dragged off to infinity”.

4. LINE ARRANGEMENTS. In this section we look at polynomials which are a product of real linear factors, and estimate the number of local maxima, minima and saddles that such polynomials can have.

The zero locus of such a polynomial is a collection of lines in the plane. We will assume that the lines are in general position, which means that they are distinct, any two have a point in common, and no three intersect at a point. We let d be the number of lines, which is the degree of the polynomial. The lines intersect in $\frac{1}{2}d(d-1)$ points. The intersection points divide the lines into d^2 segments (bounded and unbounded). Euler’s theorem applied to the plane (vertices minus edges plus faces equals one) then says that there are $\frac{1}{2}(d^2 + d + 2)$ regions (bounded and unbounded). There are $2d$ unbounded regions, so the number of bounded regions is $\frac{1}{2}(d^2 - 3d + 2) = \frac{1}{2}(d-1)(d-2)$.

Let m be the number of local maxima, n the number of local minima and s the number of saddles of the polynomial. A point at the intersection of two lines is a saddle; there are thus

$$s = \frac{1}{2}d(d-1)$$

saddles. In each region the polynomial is zero on the boundary and is either entirely positive or entirely negative in the interior. Each of the $\frac{1}{2}(d-1)(d-2)$ bounded regions must contain at least one local extremum. The polynomial thus has at least $\frac{1}{2}d(d-1) + \frac{1}{2}(d-1)(d-2) = (d-1)^2$ critical points. Since Proposition 1.2 says that this is the maximum number of critical points, the polynomial has exactly the maximum number of critical points, each critical point is nondegenerate, and there is just one extremum in each bounded region. In particular,

$$m + n = \frac{1}{2}(d-1)(d-2).$$

A convenient way to visualize a polynomial of this type is to color positive regions black and negative regions white; thus bounded black regions correspond to local maxima, and bounded white regions to local minima.

For polynomials corresponding to line arrangements it is possible to find an upper bound of the asymptotic ratio of m to n as $d \rightarrow \infty$. The following proposition appears in [Ch]; for other proofs, see the references in [FP], [To].

Proposition 4.1. *The number of local maxima m of a polynomial which is a product of d real linear factors in general position satisfies*

$$m \leq \frac{1}{3}(d^2 + d).$$

Proof: The number of local maxima m is equal to the number of bounded black regions, which is less than the number of bounded and unbounded black regions. Let ξ_i be the number of black (bounded and unbounded) regions with i sides. The total number of black regions is $\xi_2 + \xi_3 + \xi_4 + \cdots$. Now

$$3(\xi_2 + \xi_3 + \xi_4 + \cdots) \leq \xi_2 + (2\xi_2 + 3\xi_3 + 4\xi_4 + \cdots).$$

Since in an arrangement of more than two lines in general position the two-sided regions are unbounded and cannot border each other, the first term on the right, ξ_2 , is less than or equal to the number of lines d . The second term on the right is the total number of line segments d^2 . Thus the term on the right is less than or equal to $d + d^2$. \square

The exact value of this upper bound as a function of d is unknown [FP §7]. We next look at some examples of arrangements where the asymptotic upper bound of the previous proposition is attained.

Proposition 4.2. *There are arrangements of d lines in general position for which the number of local maxima and minima of the corresponding polynomial satisfy*

$$m = \frac{1}{3}d^2 + \text{terms of lower order}.$$

Chekanov [Ch] constructs the arrangements by using a lemma of Sylvester which says that it is possible to position k points in the plane so that asymptotically $k^2/6$ lines pass through exactly three points as $k \rightarrow \infty$. He also gives an ingenious proof of this lemma using real elliptic curves. The dual of Sylvester's lemma gives an arrangement of k lines for which there are asymptotically $k^2/6$ points as $k \rightarrow \infty$ at which three lines intersect. Each line is then divided into two lines very close to

one another. The regions of the resulting arrangement are now the regions of the original arrangement, the regions between each pair of close lines, and a hexagon and six triangles at each triple point of the original arrangement. The new arrangement is easily shown to have the required property.

Another construction of arrangements which satisfy Proposition 4.2 is given by Füredi and Palasti [FP]. The arrangement is constructed by choosing a particular set of d diagonals of a regular $2d$ -gon. They show that the number of triangles in this arrangement is at least $\frac{1}{3}d(d-3)$. A direct argument shows that all the triangular regions are the same color.

The results considered above on coloring an arrangement of lines can also be considered for an arrangement of circles. For circles in general position there is a conjecture [To] that the ratio of black to white regions is less than or equal to 2.5, and there is an example with four circles for which this inequality is an equality. For an arrangement of ellipses, the difficulty is to find an upper bound on the number of two-sided regions. Note that Chekanov's example can be easily converted into an arrangement of ellipses by replacing the parallel lines with ellipses that have small minor axes and large major axes, thus giving an asymptotic ratio of at least two to one for black to white regions. Coloring arrangements of curves of higher degree is the subject of the next section. It would also be interesting to investigate the problems of this section in the next higher dimension, i.e. to look at arrangements of planes in three-space.

5. HILBERT'S SIXTEENTH PROBLEM. Topologically a nonsingular projective real algebraic curve is a collection of embedded nonintersecting components each of which is homeomorphic to a circle. Hilbert's Sixteenth Problem asks about the possible arrangements of these components.

Theorem 5.1 (Harnack). *A projective real algebraic curve of degree d has at most $\frac{1}{2}(d-1)(d-2) + 1$ components.*

For several proofs of this well-known theorem, see [Gu] or [Wi]. Harnack has shown that for every d there exist curves which reach this bound. Such curves are called *M-curves*.

From now on we assume that the degree d is even. A topological component of the curve is called an *oval*. If the curve is given by the zeros in \mathbf{RP}^2 of a homogeneous polynomial $F(x, y, z)$, then an oval on whose immediate interior the function F is positive (respectively, negative) is called *even* (respectively, *odd*). (The sign of F is well defined since its degree is even.) We let

P = the number of even ovals

N = the number of odd ovals

A region to the immediate interior of an even oval must contain a local maximum of f . Thus, letting m denote the number of local maxima of f , we have

$$P \leq m.$$

The possible arrangements of even and odd ovals has been extensively examined. For a proof of the following theorem, see [Gu]:

Theorem 5.2 (Petrovskii). *For a curve of even degree d ,*

$$|P - N| \leq \frac{3}{8}d(d-2) + 1.$$

A conjecture of Ragsdale [Ra] is stronger:

Conjecture 5.3 (Ragsdale). *For a curve of even degree d*

$$P \leq \frac{3}{8}d(d-2) + 1.$$

Using methods of Hilbert and Harnack, Ragsdale gives examples of M -curves which reach the bound of 5.3. These polynomials thus have at least $(3/8)d(d-2) + 1$ maxima. Based on this result and examples of low-degree polynomials, we hesitantly make an even stronger conjecture:

Conjecture 5.4. *If $f(x, y)$ is a polynomial of even degree d with m local maxima, then*

$$m \leq \frac{3}{8}d(d-2) + 1.$$

A similar conjecture has been made by Michael Mishustin. The examples given by Ragsdale show that these bounds would be sharp. If the local maxima of f are higher than its saddles or minima, as is the case for products of lines and quadric curves in general position (Section 4), then by taking a level curve just below the maxima and above the saddles and minima we have $P = m$ and $N = 0$, and 5.2 implies 5.4. Polynomials of odd degree appear to have similar properties.

6. POLYNOMIALS OF INDEX GREATER THAN ONE. In this section we show that a polynomial with index greater than one must have critical points at infinity. These results are due to Jeff Roy. Let (x, y) be coordinates for points in the plane \mathbf{R}^2 as before, and let $[x, y, z]$ be homogeneous coordinates for points in real projective space \mathbf{RP}^2 . We define the *line at infinity*, written as L_∞ , to be $\{z = 0\}$.

Definition 6.1. A point $p \in L_\infty$ is a *critical point at infinity* for the polynomial f if the projective completions of the curves $f_x = 0$ and $f_y = 0$ intersect at p .

For example, the polynomial $f(x, y) = xy^2 - y$ has a critical point at $[1, 0, 0]$. (This definition of ‘critical point at infinity’ is not the usual one and is only for the purposes of this paper.) The main result of this section is as follows:

Theorem 6.2. *A polynomial with isolated critical points and index $i > 1$ has at least $2(i-1)$ critical points at infinity (counted with multiplicities).*

The proof of 6.2 relies on the fact that the index of the gradient vector field is determined by the behavior of the gradient near infinity. In particular, given a polynomial f of degree d with index i , on a sufficiently large circle there must be at least $i-1$ points where the gradient points directly towards the origin, and at least $i-1$ points where the gradient points directly away from the origin. The points where the gradient points directly towards or away from the origin are given by an algebraic curve. We show that this curve has unbounded components, and that $f(p)$ approaches a finite limit as p goes to infinity along each component. We then show that these components are tangent at infinity and that f has multiple critical points at the points of tangency. The proof of Theorem 6.2 will occupy the rest of this section. We begin with a few lemmas on algebraic curves.

Definition 6.3. A smooth point p_0 on an algebraic curve G is a *turning point* relative to a point $p \in \mathbf{R}^2$ if the tangent to G at p_0 is perpendicular to the line from p to p_0 .

Lemma 6.4. *If G is a reduced curve and p is a point not on G , there is an annulus centered at p containing all singularities and turning points of G relative to p .*

Proof: Let $G = \{g(x, y) = 0\}$ and without loss of generality let p be the origin. Recall that G is reduced if the polynomial g is square-free. First we consider the turning points. The tangent to G at a smooth point (x_0, y_0) is given by

$$(x - x_0)g_x(x_0, y_0) + (y - y_0)g_y(x_0, y_0) = 0.$$

If this tangent is perpendicular to the line from the origin to (x_0, y_0) , we get $y_0g_x(x_0, y_0) - x_0g_y(x_0, y_0) = 0$. Let

$$h(x, y) = yg_x - xg_y$$

and let $H = \{h(x, y) = 0\}$. The intersection of the curves G and H contains the turning points of G . (If $h = 0$, then H is the whole plane and every point of G is a turning point.) A common component of G and H is an arc of a circle centered at the origin; this follows from the uniqueness theorem for ordinary differential equations since at each smooth point the tangent line is perpendicular to the line through the origin. By Bezout's Theorem the remaining set of intersections of G and H forms a finite set.

Since G is reduced it has (a finite number of) isolated singular points. Thus an annulus can be chosen containing the turning points and singularities of G . \square

Lemma 6.5. *Let v be a vector field in the plane and let C be a circle centered at the origin. If v does not vanish on C and has index $i > 1$ around C , then there exist $2(i - 1)$ arcs on C such that for each arc (p_0, p_1)*

1. *v points directly away from the origin at p_0 and directly towards the origin at p_1 , and*
2. *at all points on the arc (p_0, p_1) , the component of v tangent to C is positive in the direction from p_0 to p_1 .*

Proof: Without loss of generality assume that C is the unit circle S^1 . Define a new vector field $w: S^1 \rightarrow S^1$ by

$$w(z) = \frac{v(z)}{\|v(z)\|} z^{-1}.$$

The index of $v(z)/\|v(z)\|$ is i and the index of the vector field $z \mapsto z^{-1}$ is -1 , so the index of the product $w(z)$ is $i - 1$. The vector field w maps all points where v points directly away from the origin to 1, and maps all points where v points directly towards the origin to -1 .

Let A be the upper half circle from 1 to -1 . The set $w^{-1}(A)$ will be a set of arcs in S^1 . As the index of w is $i - 1$, at least $i - 1$ of these arcs must be of the form (p_0, p_1) , where $w(p_0) = 1$, $w(p_1) = -1$, and $w^{-1}(p)$ moves in the positive direction from 1 to -1 as p moves from p_0 to p_1 . Thus for each of these arcs we have $v(p_0)$ points directly away from 0, $v(p_1)$ points directly towards 0, and the component of v tangent to S^1 is positive in the direction from p_0 to p_1 at each point on the arc.

Letting A' be the bottom half circle from -1 to 1, we obtain another $i - 1$ arcs in $w^{-1}(A')$ with these properties. This gives us a total of $2(i - 1)$ arcs, as required. \square

Now suppose that $f(x, y)$ is a polynomial with isolated critical points and with index $i > 1$. The gradient field of the polynomial forms a vector field in \mathbf{R}^2 to which we can apply Lemma 6.5. The points where $\text{grad}(f)$ vanishes or points straight away from or straight towards the origin are given by the curve G defined by

$$G = \{yf_x - xf_y = 0\}.$$

By Lemma 6.4 there is an annulus containing all the singular points and turning points of the reduced curve corresponding to G . We choose the outer circle C of the annulus large enough to contain all critical points of f . We then apply Lemma 6.5 to C ; for each arc the endpoints p_0 and p_1 are where C intersects G .

Lemma 6.6. *For each arc of Lemma 6.5 the curve G has unbounded topological components G_0 and G_1 in the exterior of the circle C above with parameterizations $G_0(t)$ and $G_1(t)$ for $0 \leq t < \infty$ such that $G_0(0) = p_0$, $G_1(0) = p_1$, $\lim_{t \rightarrow \infty} G_0(t) = \lim_{t \rightarrow \infty} G_1(t) = \infty$, $\lim_{t \rightarrow \infty} f(G_0(t)) = \alpha_0$, and $\lim_{t \rightarrow \infty} f(G_1(t)) = \alpha_1$, where α_0 and α_1 are finite constants with $\alpha_0 \leq \alpha_1$.*

Proof: (See Figure 6.1.) By Lemma 6.5 the tangential component of $\text{grad}(f)$ is positive in the direction from p_0 towards p_1 , implying

$$f(p_1) > f(p_0).$$

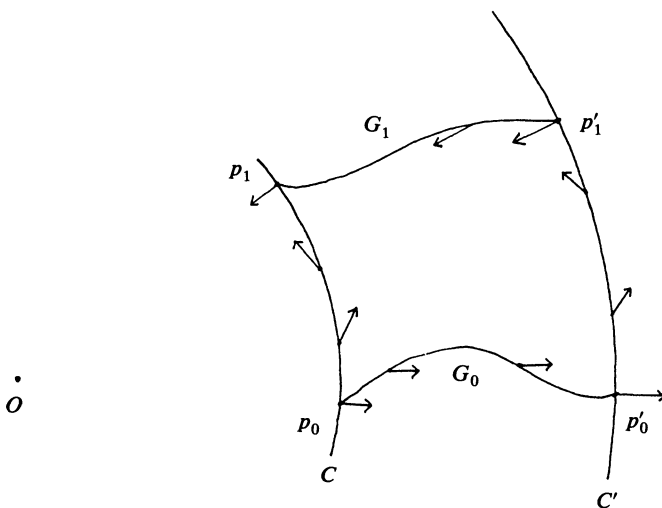


Figure 6.1

Consider the topological component G_0 of G beginning at p_0 and lying exterior to C . As all singularities and turning points of the reduced curve corresponding to G lie within C , the component G_0 is unbounded. Parameterize G_0 so that $G_0(0) = p_0$ and $G_0(t) \rightarrow \infty$ as $t \rightarrow \infty$. Likewise, parameterize the component G_1 of G beginning at p_1 so that $G_1(0) = p_1$ and $G_1(t) \rightarrow \infty$ as $t \rightarrow \infty$.

As all critical points of f are contained within C , the vector field $\text{grad}(f)$ is nonvanishing on both G_0 and G_1 . Since all turning points of G are contained within C , the component of $\text{grad}(f)$ tangent to G_0 is positive in the direction of

increasing t , and the component tangent to G_1 is negative in the direction of increasing t . Thus as t increases, $f(G_0(t))$ monotonically increases and $f(G_1(t))$ monotonically decreases.

For any point p'_0 on G_0 , let C' be the circle centered at the origin which intersects G_0 at p'_0 . Let p'_1 be the point of intersection of C' and G_1 . By the above,

$$f(p'_0) > f(p_0)$$

and

$$f(p_1) > f(p'_1).$$

Since we are beyond all turning points and singular points of G , the curve G does not intersect C' anywhere on the arc (p'_0, p'_1) . Thus, the tangential component of $\text{grad}(f)$ is either always positive or always negative from p'_0 to p'_1 . If it is always negative, then the index of $\text{grad}(f)$ around the loop (p_0, p'_0, p'_1, p_1) is -1 , but this is impossible, as all critical points of f are contained in C . Therefore the tangential component of $\text{grad}(f)$ points from p'_0 to p'_1 , implying

$$f(p'_1) > f(p'_0).$$

The last two inequalities imply that f is bounded on G_0 . Since f is increasing, there is a constant α_0 such that f approaches α_0 along G_0 . A similar argument shows that f approaches a finite limit α_1 along G_1 . Clearly $\alpha_0 \leq \alpha_1$. \square

In fact, all our examples have $\alpha_0 = \alpha_1$. A similar argument applied to the inner circle of the annulus of 6.4 gives an alternate proof of 2.6.

Lemma 6.7. *Under the hypotheses of Lemma 6.6, G_0 and G_1 are asymptotic, i.e. tangent at a point on the line at infinity in \mathbf{RP}^2 .*

Proof: Let q be the point at which G_0 meets the line at infinity. Assume that G_0 is not tangent to G_1 at q . Then there exists an infinite number of lines which pass through q and lie between G_0 and G_1 near the line at infinity. Let l be one such line, and assume that l lies between G_0 and G_1 beyond some circle centered at the origin. Let C' be any circle centered at the origin larger than this one. As in the proof of Lemma 6.6, assume that C' intersects G_0 and G_1 at p'_0 and p'_1 respectively. We know for all p' on the arc (p'_0, p'_1) that

$$f(p_0) < f(p'_0) < f(p') < f(p'_1) < f(p_1).$$

Since this is true for all circles C' , the function f is bounded on this end of the line l . Since there is an infinity of such lines, the proof of Proposition 1.3 implies that f is a polynomial in one variable after a change of coordinates. Such a polynomial has nonisolated critical points since by Proposition 2.5, $i > 1$ implies that $d > 2$. Hence an infinite number of such lines cannot exist, and G_0 and G_1 are tangent at q . \square

We are now ready to prove the main theorem of this section.

Proof of Theorem 6.2. Starting with a polynomial $f(x, y)$ with index $i > 1$, we have around a large circle C the $2(i - 1)$ arcs described by Lemma 6.5. For each arc, we have found topological components G_0 and G_1 of the curve G which are tangent at a point q on the line at infinity. Since the gradient vector field winds more than 180 degrees as we move on the arc of Lemma 6.5 from the point p_0 to the point

p_1 , the curves given by $f_x = 0$ and $f_y = 0$ must intersect that arc. By choosing the circle C of Lemma 6.5 sufficiently large so that all singularities and turning points of the curves $f_x = 0$ and $f_y = 0$ are inside C , there must be topological components of $f_x = 0$ and $f_y = 0$ extending to infinity and bounded by the components G_0 and G_1 . Thus $f_x = 0$ and $f_y = 0$ are likewise tangent at q and hence have intersection multiplicity of at least two there.

Each of the $2(i - 1)$ arcs thus contributes two critical points at infinity. However, since two arcs may contribute points at infinity on the opposite side of the real plane and thus the same point in the real projective plane, there are at least $2(i - 1)$ critical points at infinity. \square

Remark 6.8. In all examples the critical point p at infinity has ‘finite critical value’ c in the following sense [Ha]: There is a sequence in \mathbf{R}^2 of points $p_k \rightarrow p$ such that $f(p_k) \rightarrow c$ and $\text{grad } f(p_k) \rightarrow 0$.

Next we give an improved form of Inequality 3.2.

Corollary 6.9. *If $f(x, y)$ is a polynomial of degree d with m maxima and n minima, then*

$$m + n \leq \frac{1}{2}d^2 - d + 1.$$

Proof: If $i \leq 1$, Proposition 1.2 implies $m + n + s \leq (d - 1)^2$, and by the index theorem $m + n - s \leq 1$. Adding these we get $2(m + n) \leq d^2 - 2d + 2$, and thus $m + n \leq \frac{1}{2}d^2 - d + 1$. If $i > 1$, Proposition 1.2 (which bounds the number of critical points at infinity as well) implies $m + n + s \leq (d - 1)^2 - 2(i - 1)$, and by the index theorem $m + n - s = i$. Adding these we get $2(m + n) \leq d^2 - 2d + 3 - i < d^2 - 2d + 2$ and thus $m + n < \frac{1}{2}d^2 - d + 1$. This completes the proof. \square .

In all of the examples we have looked at, however, the critical points constructed above have multiplicity much higher than two, since the components of $f_x = 0$ and $f_y = 0$ which are tangent at infinity have high multiplicity. Our examples support the following conjecture:

Conjecture 6.10. *If $f(x, y)$ is a polynomial with index $i > 1$, then f has at least $12(i - 1)$ critical points at infinity.*

All calculus students know (or should know) that a function of one variable with a local maximum at a point and no other critical points has an absolute maximum at that point. For functions of two variables the analogous result is not true: There are functions with a local maximum and no other critical points, but this critical point is not an absolute maximum. Nice pictures of such functions can be found in [Sm] and [Wag, Sect. 3.4 and Ch. 7]. For example [CV], the polynomial

$$x^2(1 + y)^3 + y^2$$

has this property; the origin is the only critical point. One would expect, in analogue with Theorem 6.2, that such polynomials would have a critical point at infinity, and in fact it is not hard to show that this is true [Du]. (A critical point at infinity can clearly be seen in the pictures.)

7. REMARKS. It would be interesting to investigate in greater detail the asymptotic behavior of the number of local maxima of real polynomials with isolated critical points as the degree becomes large. For polynomials of one variable

$$m \leq \frac{1}{2}d + \text{constant}.$$

For polynomials of two variables which are a product of real linear factors, Proposition 4.1 says that

$$m \leq \frac{1}{3}d^2 + \text{terms of lower order}.$$

Examples attaining these bounds were given in that section. For arrangements of circles the conjecture at the end of Section 5 combined with 3.4 would imply that

$$m \leq \frac{5}{14}d^2 + \text{terms of lower order}$$

and for arbitrary polynomials of two variables Conjecture 5.4 says that

$$m \leq \frac{3}{8}d^2 + \text{terms of lower order}.$$

Problem 7.1. *What are these asymptotic estimates for polynomials of three or more variables?*

The methods of this paper are equally valid for nowhere-vanishing rational functions, i.e. functions of the form $f(x, y)/g(x, y)$ where $f(x, y)$ and $g(x, y)$ are real polynomials, and $g(x, y) \neq 0$ for all points $(x, y) \in \mathbf{R}^2$. Note that rational functions do not satisfy Proposition 1.3; for example, the rational function $1/(x^2 + y^2)$ is bounded on the whole plane, and in fact has its zero level curve ‘at infinity’.

The problems of this paper can also be considered in the local case, by analyzing the types of critical points of a nearby Morse function to a function with an isolated critical point [Ar2, Problem 5]. This case is probably quite different from the global case (c.f. Proposition 2.6).

REFERENCES

-
- [Ar1] V. I. Arnold, Index of a Singular Point of a Vector Field, the Petrovskii-Oleinik Inequality, and Mixed Hodge Structures, *Functional Analysis and its Applications* 12 (1978) 1–11.
 - [Ar2] ———, On some problems in singularity theory. In: *Geometry and analysis, papers dedicated to the memory of V. K. Patodi*, Indian Academy of Sci. p. 1–9.
 - [BK] E. Brieskorn and H. Knörrer, *Plane Algebraic Curves*, Birkhäuser Verlag, Boston, 1986.
 - [CV] B. Calvert and M. K. Vamanamurthy, Local and Global Extrema for Functions of Several Variables, *J. Austral. Math. Soc. (Series A)*, 29 (1980) 362–368.
 - [Ch] V. Chekanov, Asymptotics of the Number of Maxima for a Product of Linear Functions of Two Variables, *Moscow University Mathematics Bulletin* 41 (1986) 85–87.
 - [Da] R. Davies, Problems, *Mathematics Magazine* 61 (1988) 59.
 - [Du] A. Durfee, Critical points at infinity (in preparation).
 - [Fu] W. Fulton, *Algebraic Curves*, Benjamin 1969.
 - [FP] Z. Füredi and I. Palasti, Arrangements of Lines with a Large Number of Triangles, *Proceedings of the American Mathematical Society* 92 (1984) 561–566.
 - [GP] V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall 1974.
 - [Gri] P. Griffiths, *Introduction to Algebraic Curves*, American Mathematical Society, 1989.
 - [Grü] B. Grünbaum, *Arrangements and Spreads*, Amer. Math. Soc., Providence, 1965.
 - [Gu] D. A. Gudkov, The Topology of Real Projective Algebraic Varieties, *Russian Math Surveys* 29:4 (1974) 1–79.
 - [Ha] Ha Huy Vui, Number of Lojasiewicz and singularities at infinity of polynomial functions of two variables, *Math. Inst. Hanoi* Preprint 1990.
 - [Mi] J. Milnor, *Topology from the Differentiable Viewpoint*, The University Press of Virginia, Charlottesville, 1965.

- [Ra] V. Ragsdale, On the Arrangement of the Real Branches of Plane Algebraic Curves, *American J. Math.* 28 (1906) 377–404.
- [Ro] I. Rosenholtz, Problem 1235, *Mathematics Magazine* 59 (1986) 44.
- [Sm] D. Smith, ed., Notes, *Mathematics Magazine* 58 (1985) 146–150.
- [To] L. Tóth, A combinatorial problem concerning oriented lines in the plane, *Amer. Math. Monthly* 82 (1975) 387–389.
- [Wag] S. Wagon, *Mathematica in Action*, W. H. Freeman and Co., New York, 1991.
- [Wal] R. Walker, *Algebraic Curves*, Princeton University Press, 1950.
- [Wi] G. Wilson, Hilbert's Sixteenth Problem, *Topology* 17 (1978) 53–73.

*Department of Mathematics
Mount Holyoke College
South Hadley, MA 01075*

I like to look at mathematics almost more as an art than as a science; for the activity of the mathematician, constantly creating as he is, guided though not controlled by the external world of the senses, bears a resemblance, not fanciful I believe but real, to the activity of an artist, of a painter let us say. Rigorous deductive reasoning on the part of the mathematician may be likened here to technical skill in drawing on the part of the painter. Just as no one can become a good painter without a certain amount of skill, so no one can become a mathematician without the power to reason accurately up to a certain point. Yet these qualities, fundamental though they are, do not make a painter or mathematician worthy of the name, no indeed are they the most important factors in the case. Other qualities of a far more subtle sort, chief among which in both cases is imagination, go to the making of a good artist or good mathematician.

—*Maxime Bôcher*

NOTES

Edited by: John Duncan

Separation of the Zeros of Polynomials

Peter Walker

Let $f(x) = \prod_1^n (x - a_i)$ be a polynomial of degree n with distinct real zeros which we shall assume are in increasing order: $a_1 < a_2 < \cdots < a_n$. Rolle's theorem tells us that there is exactly one zero of f' in each interval (a_i, a_{i+1}) , and intuitively we feel that the zeros of f' are more widely or more evenly distributed than those of f . To support this idea, we shall measure the separation of the zeros by $\delta(f) = \min_i (a_{i+1} - a_i)$ and show that $\delta(f') > \delta(f)$.

More generally it is easy to see by considering the graph of f'/f that for given real k , the polynomial $f' - kf$ has one zero, which we shall denote by b_i , in each (a_i, a_{i+1}) . In addition if $k > 0$ there is one further zero $b_n > a_n$, while if $k < 0$ there is one further zero $b_0 < a_1$. In this case we also have $\delta(f' - kf) > \delta(f)$.

A deeper analysis shows that in fact the ratio $\delta(f' - kf)/\delta(f)$ is bounded away from 1; in particular there is a constant $c_n > 1$ such that $\delta(f') \geq c_n \delta(f)$. These results will appear elsewhere.

Theorem. $\delta(f' - kf) > \delta(f)$.

Proof: Let $d = \delta(f)$; we have to show that $b_j - b_{j-1} > d$ for all j .

Let

$$g(x) = \frac{f'(x)}{f(x)} = \sum_1^n \frac{1}{x - a_i};$$

we have $g(b_j) = g(b_{j-1}) = k$, and so

$$\sum_{i=1}^n \frac{1}{(b_{j-1} - a_i)(b_j - a_i)} = 0$$

since $b_{j-1} \neq b_j$.

Fix a value of j , and write $b_{j-1} = a_j - u$ and $b_j = a_j + v$. Then for $i \geq 1$ put $D_i = a_j - a_{j-i}$, and $E_i = a_{j+i} - a_j$, so that the above can be written

$$\sum_{i=1}^{j-1} \frac{1}{(D_i - u)(D_i + v)} + \frac{1}{(-u)(v)} + \sum_{i=1}^{n-j} \frac{1}{(E_i + u)(E_i - v)} = 0,$$

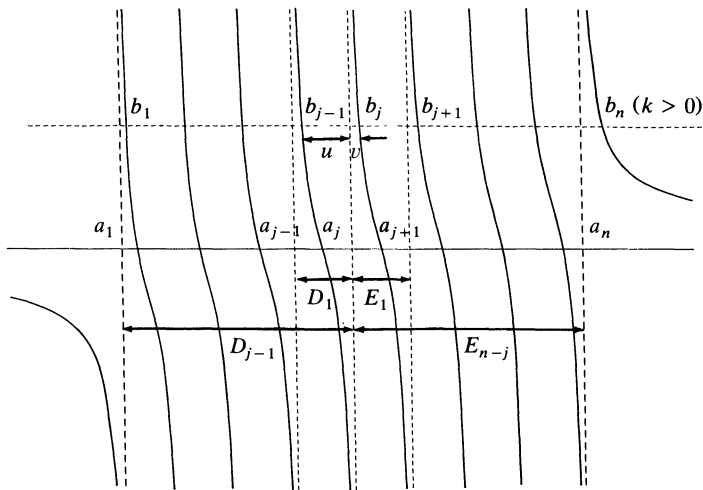


Figure 1

or equivalently

$$F(u, v) = \sum_{i=1}^{j-1} \frac{uv}{(D_i - u)(D_i + v)} + \sum_{i=1}^{n-j} \frac{uv}{(E_i + u)(E_i - v)} = 1.$$

Now the function F is defined and continuous for $0 \leq u < D_1$ and $0 \leq v < E_1$, and is strictly increasing in each variable separately. We want to show that if $F(u, v) = 1$ then $b_j - b_{j-1} = u + v > d$; equivalently we shall show that if $u + v = d$ then $F(u, v) < 1$.

But for $0 < u < d$, $v = d - u$ we have

$$F(u, d - u) = u(d - u) \left[\sum_{i=1}^{j-1} \frac{1}{(D_i - u)(D_i + d - u)} + \sum_{i=1}^{n-j} \frac{1}{(E_i + u)(E_i - d + u)} \right],$$

and since both $D_i, E_i \geq id$ we obtain

$$\begin{aligned} F(u, d - u) &\leq u(d - u) \left[\sum_{i=1}^{j-1} \frac{1}{(id - u)(id + d - u)} + \sum_{i=1}^{n-j} \frac{1}{(id + u)(id - d + u)} \right] \\ &< \frac{u(d - u)}{d} \left[\frac{1}{d - u} + \frac{1}{u} \right] = 1, \end{aligned}$$

since both sums telescope; the result follows.

College of Science
Sultan Qaboos University
P.O. Box 32486
Al-Khod, Muscat
Sultanate of Oman

A Short Proof of Jacobi's Formula for the Number of Representations of an Integer as a Sum of Four Squares

George E. Andrews, Shalosh B. Ekhad and Doron Zeilberger

Diophantus probably knew, and Lagrange [L] proved, that every positive integer can be written as a sum of four perfect squares. Jacobi [J] proved the stronger result that the number of ways in which a positive integer can be so written¹ equals 8 times the sum of its divisors that are not multiples of 4. Here we give a short new proof that only uses high school algebra, and is completely *from scratch*. All infinite series and products that appear are to be taken in the entirely elementary sense of formal power series.

The problem of representing integers as sums of squares has drawn the attention of many great mathematicians, and we encourage the reader to look up Grosswald's [G] erudite masterpiece on this subject.

The crucial part of our proof is played by two simple identities, that we state as one Lemma.

Lemma. *Let*

$$H_n = H_n(q) = \frac{1+q}{1-q} \frac{1+q^2}{1-q^2} \cdots \frac{1+q^n}{1-q^n}.$$

For all integers $n \geq 0$,

$$\sum_{k=-n}^n \frac{4(-q)^k}{(1+q^k)^2} H_n^2 H_{n+k} H_{n-k} = 1, \quad (a)$$

$$\sum_{k=0}^n \frac{2(-q^{n+1})^k}{1+q^k} \frac{H_k}{H_n} = \sum_{k=-n}^n (-q)^{k^2}. \quad (b)$$

Proof: Let $L_1(n)$ and $L_2(n)$ be the left sides of (a) and (b) respectively, and let $F_1(n, k)$, and $F_2(n, k)$ be the respective summands. Since both (a) and (b) obviously hold for $n = 0$, it suffices to prove that for every $n \geq 0$, $L_1(n+1) - L_1(n) = 0$, and $L_2(n+1) - L_2(n) = 2(-q)^{(n+1)^2}$. To this end, we construct

$$G_1(n, k) := \frac{q^{n-k+1}(1+q^{2n+2})(1+q^k)^2(1+q^{n+k+1})}{(1-q^{n+1})^3(1-q^{n+k+1})(1+q^{n+1})} F_1(n, k),$$

$$G_2(n, k) := \frac{(-q^{n+1})(1+q^k)}{1+q^{n+1}} F_2(n, k),$$

¹Precisely: the number of vectors (*not sets*) (x_1, x_2, x_3, x_4) , where the components are (positive, negative, or zero) integers, such that $x_1^2 + x_2^2 + x_3^2 + x_4^2 = n$.

with the motive that

$$\begin{aligned} F_1(n+1, k) - F_1(n, k) &= G_1(n, k) - G_1(n, k-1), \\ F_2(n+1, k) - F_2(n, k) &= G_2(n, k) - G_2(n, k-1), \end{aligned} \quad (1)$$

which immediately imply them, by telescoping, upon summing from $k = -n - 1$ to $k = n + 1$, and from $k = 0$ to $k = n + 1$ respectively. The two identities of (1) are purely routine, since dividing through by $F_1(n, k)$ and $F_2(n, k)$ respectively, lead to routinely-verifiable high-school-algebra identities. \square

Dividing both sides of (a) by H_n^4 and letting $n \rightarrow \infty$ in (a) and (b) gives

$$1 + 8 \sum_{k=1}^{\infty} \frac{(-q)^k}{(1+q^k)^2} = H_{\infty}^{-4}, \quad (a')$$

$$\sum_{k=-\infty}^{\infty} (-q)^{k^2} = H_{\infty}^{-1}. \quad (b')$$

Combining (a') and (b'), yields, after changing $q \rightarrow -q$,

$$\left(\sum_{k=-\infty}^{\infty} q^{k^2} \right)^4 = 1 + 8 \sum_{k=1}^{\infty} \frac{q^k}{(1+(-q)^k)^2}. \quad (2)$$

The coefficient of a typical term q^n on the left of (2) is the number of ways of writing n as a sum of four squares. It remains to show that the coefficient of q^n in the sum on the right of (2) equals the sum of the divisors of n that are not multiples of 4.

Using the power-series expansion $z/(1+z)^2 = \sum_{r=1}^{\infty} (-1)^{(r+1)} r z^r$, with $z = (-q)^k$, and collecting like powers, the sum on the right side may be rewritten

$$\sum_{k=1}^{\infty} \sum_{r=1}^{\infty} (-1)^{(k+1)(r+1)} r q^{kr} = \sum_{n=1}^{\infty} q^n \left[\sum_{r|n} (-1)^{(r+1)(n/r+1)} r \right].$$

The coefficient of q^n above is a weighed sum of divisors r of n , where the coefficient of r is $-1 = +1 - 2$ if both r and n/r are even and $+1$ otherwise, so the coefficient of q^n is

$$\sum_{r|n} r - \sum_{\substack{r|n \\ r, n/r \text{ even}}} 2r = \sum_{d|n} d - \sum_{\substack{d|n \\ 4|d}} d = \sum_{\substack{d|n \\ 4 \nmid d}} d. \quad \square$$

The finitary identities (a) and (b) combine to yield a single finitary identity

$$\left(\sum_{k=0}^n \frac{2(-q^{n+1})^k}{1+q^k} H_k \right)^4 \sum_{k=-n}^n \frac{4(-q)^k}{(1+q^k)^2} \frac{H_{n+k}}{H_n} \frac{H_{n-k}}{H_n} = \left(\sum_{k=-n}^n (-q)^{k^2} \right)^4, \quad (3)$$

which also immediately implies Jacobi's theorem, by taking it "mod q^n " for any desired n . Identity (3) makes it transparent that our proof only uses the potential infinity, not the ultimate one.

The identities of the Lemma are examples of q -binomial coefficient identities, a.k.a terminating basic hypergeometric series identities. The proof of such identities is now completely routine [WZ] [Z]. The proof of the Lemma given here used the algorithm of [Z]. Further applications of basic hypergeometric series to number theory can be found in [A1]. An excellent modern reference to basic hypergeometric series is [GR].

We conclude with some comments addressed mainly to the cognoscenti. Identities (a) and (b) are special cases of classical identities: (a) is a special case of

Jackson's theorem [GR, p. 35, eq. (2.6.2)], and (b) is a special case of Watson's q -analog of Whipple's theorem ([GR, p. 35, eq. (2.5.1)], see also [A2, p. 118, eq. (4.3)].) The discovery of (b) was motivated by [S1] and [S2].

We see fairly clearly how to do the 2-square theorem (a different instance of Jackson's theorem replaces (a)); however the theorems for 6 and 8 squares apparently require (using this approach) some instance of the ${}_6\Psi_6$ summation theorem [GR, p. 128, (5.3.1)] (see [A1, pp. 461–465] for details). Since we do not know a finitary analog of the ${}_6\Psi_6$ summation, the question of a similar result for 6 and 8 squares is of interest.

ACKNOWLEDGMENT. The referee made several helpful comments.

REFERENCES

-
- [A1] G. E. Andrews, Applications of basic hypergeometric functions, *SIAM Rev.* 16 (1974), 441–484.
 - [A2] G. E. Andrews, The fifth and seventh order mock theta functions, *Trans. Amer. Math. Soc.*, 293 (1986), 113–134.
 - [GR] G. Gasper and M. Rahman, *Basic hypergeometric series*, Cambridge University Press, 1990.
 - [G] E. Grosswald, *Representations of integers as sums of squares*, Springer, New York, 1985.
 - [J] C. G. J. Jacobi, Note sur la décomposition d'un nombre donné en quatre carrés, *J. Reine Angew. Math.* 3 (1828), 191. *Werke*, vol. I, 247.
 - [L] J. L. Lagrange, Nouveau Mém. Acad. Roy. Sci. Berlin (1772), 123–133; *Oeuvres*, vol. 3, 189–201.
 - [S1] D. Shanks, A short proof of an identity of Euler, *Proc. Amer. Math. Soc.*, 2 (1951), 747–749.
 - [S2] D. Shanks, Two theorems of Gauss, *Pacific. J. Math.*, 8 (1958), 609–612.
 - [WZ] H. S. Wilf and D. Zeilberger, An algorithmic proof theory for multisum/integral (ordinary and “ q ”) hypergeometric identities, *Invent. Math.* 108 (1992), 575–633.
 - [Z] D. Zeilberger, The method of creative telescoping for q -series, in preparation.

Andrews:

Department of Mathematics

Pennsylvania State University

University Park, PA 16182

andrews@cantor.math.psu.edu

Ekhad & Zeilberger:

Department of Mathematics

Temple University

Philadelphia, PA 19122

ekhad@euclid.math.temple.edu

zeilberg@euclid.math.temple.edu

An Elementary Proof of Hilbert's Inequality

Krzysztof Oleszkiewicz

SUMMARY. We present an elementary proof of Hilbert's inequality (1) and we give a simple example showing that the constant in (1) is optimal. Moreover, we give a (slightly less elementary) extension of (1). There are a lot of applications of Hilbert's inequality to the theory of analytic functions and to the theory of functions of a real variable; some of them can be found in Chapter IX and Appendix 3 of Hardy, Littlewood and Polya's *Inequalities*, Cambridge 1964. The original proof of the inequality, due to Hilbert, is also given in that book.

Proposition 1 (Hilbert's inequality). *If (a_m) and (b_n) are square summable sequences of real numbers, then the double series $\sum_{m,n=1}^{\infty} a_m b_n / (m + n)$ is convergent*

and

$$\sum_{m,n=1}^{\infty} \frac{a_m b_n}{m+n} \leq \pi \sqrt{\sum_{m=1}^{\infty} a_m^2} \cdot \sqrt{\sum_{n=1}^{\infty} b_n^2}. \quad (1)$$

The inequality is strict unless one of the sequences (a_m) or (b_n) is identically zero. Moreover, π is the best constant in (1).

Proof of Proposition 1. We will use two lemmas.

Lemma 1. For each positive number m , $\sum_{n=1}^{\infty} (\sqrt{m}/\sqrt{n}(m+n)) < \pi$.

Proof: Let us denote points $(0, 0)$, $(0, \sqrt{m})$, (\sqrt{m}, \sqrt{n}) by C, Y, X_n ($n = 0, 1, 2, \dots$) respectively. Let S be the area of the quadrant of the circle centred at C with radius \sqrt{m} from X_0 to Y . Let us denote by R_n the intersection of the circle and the line CX_n . Then let B_n be the intersection of the line CX_{n-1} and the vertical line containing the point R_n (for $n = 1, 2, 3, \dots$). Moreover, let S_n denote the area of the sector $R_{n-1}CR_n$ of the circle. (See picture 1.)

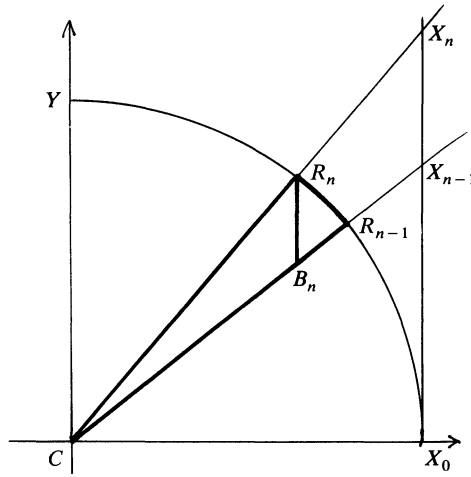


Figure 1

Denoting the area of the triangle KLM by $S_{\Delta KLM}$ we find

$$\begin{aligned} \frac{\pi m}{4} = S &= \sum_{n=1}^{\infty} S_n > \sum_{n=1}^{\infty} S_{\Delta R_n C B_n} \\ &= \sum_{n=1}^{\infty} \left(\frac{|CR_n|}{|CX_n|} \right)^2 S_{\Delta X_{n-1} C X_n} \\ &= \sum_{n=1}^{\infty} \frac{m}{|CX_0|^2 + |X_0 X_n|^2} \cdot \frac{|CX_0| \cdot |X_{n-1} X_n|}{2} \\ &= \sum_{n=1}^{\infty} \frac{m\sqrt{m}(\sqrt{n} - \sqrt{n-1})}{2(m+n)} \\ &> \sum_{n=1}^{\infty} \frac{m\sqrt{m}}{4\sqrt{n}(m+n)} \end{aligned}$$

and therefore

$$\sum_{n=1}^{\infty} \frac{\sqrt{m}}{\sqrt{n}(m+n)} < \pi. \quad \square$$

The generalization of Lemma 1 will be given later, in Lemma 3. Now we will prove Hilbert's inequality. Writing

$$\frac{a_m b_n}{m+n} = \frac{\sqrt[4]{m}}{\sqrt{n}\sqrt{m+n}} a_m \cdot \frac{\sqrt[4]{n}}{\sqrt[4]{m}\sqrt{m+n}} b_n$$

and using Schwarz's inequality we find

$$\begin{aligned} \sum_{m,n=1}^{\infty} \frac{a_m b_n}{m+n} &\leq \sqrt{\sum_{m,n=1}^{\infty} \frac{\sqrt{m}}{\sqrt{n}(m+n)} a_m^2} \cdot \sqrt{\sum_{m,n=1}^{\infty} \frac{\sqrt{n}}{\sqrt{m}(m+n)} b_n^2} \\ &= \sqrt{\sum_{m=1}^{\infty} \left(\sum_{n=1}^{\infty} \frac{\sqrt{m}}{\sqrt{n}(m+n)} \right) a_m^2} \cdot \sqrt{\sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} \frac{\sqrt{n}}{\sqrt{m}(m+n)} \right) b_n^2} \\ &\leq \pi \sqrt{\sum_{m=1}^{\infty} a_m^2} \cdot \sqrt{\sum_{n=1}^{\infty} b_n^2}, \end{aligned}$$

where Lemma 1 was used. Obviously the last inequality is strict unless one of the sequences (a_m) or (b_n) is identically zero. Now we will prove that π cannot be replaced by any smaller constant.

Lemma 2. For each natural number $m > 1$

$$\sum_{n=1}^{m-1} \frac{1}{\sqrt{mn}(m+n)} > \frac{\pi}{2m} - \frac{2}{m\sqrt{m}}.$$

Proof: Let us denote by A_n the intersection of the line CX_{n+1} and the line $R_n B_n$ (for $n = 0, 1, \dots, m-1$). Let S' be the area of the sector $X_0 C X_m$ of the circle. (See picture 2.) Then clearly

$$\begin{aligned} \frac{\pi m}{8} = S' &< \sum_{n=0}^{m-1} S_{\Delta R_n C A_n} = \frac{\sqrt{m}}{2} + \sum_{n=1}^{m-1} \left(\frac{|CR_n|}{|CX_n|} \right)^2 S_{\Delta X_n C X_{n+1}} \\ &= \frac{\sqrt{m}}{2} + \sum_{n=1}^{m-1} \frac{m\sqrt{m}|X_n X_{n+1}|}{2(|CX_0|^2 + |X_0 X_n|^2)} \\ &= \frac{\sqrt{m}}{2} + \sum_{n=1}^{m-1} \frac{m\sqrt{m}(\sqrt{n+1} - \sqrt{n})}{2(m+n)} \\ &< \frac{\sqrt{m}}{2} + \sum_{n=1}^{m-1} \frac{m\sqrt{m}}{4\sqrt{n}(m+n)}. \end{aligned}$$

Thus

$$\sum_{n=1}^{m-1} \frac{1}{\sqrt{mn}(m+n)} > \frac{\pi}{2m} - \frac{2}{m\sqrt{m}}. \quad \square$$

Proof: We can estimate

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{m^{1/p}}{n^{1/p}(m+n)} &\leq \int_0^{\infty} \frac{m^{1/p} dx}{x^{1/p}(x+m)} = \int_0^{\infty} \frac{dt}{t^{1/p}(1+t)} \\
&= \int_0^1 \frac{dt}{t^{1/p}(1+t)} + \int_1^{\infty} \frac{dt}{t^{1+1/p}\left(1+\frac{1}{t}\right)} \\
&= \int_0^1 \sum_{n=0}^{\infty} (-1)^n t^{n-1/p} dt + \int_1^{\infty} \sum_{n=0}^{\infty} (-1)^n t^{-n-1-1/p} dt \\
&= \sum_{n=0}^{\infty} (-1)^n \int_0^1 t^{n-1/p} dt + \sum_{n=0}^{\infty} (-1)^n \int_1^{\infty} t^{-n-1-1/p} dt \\
&= \sum_{n=1}^{\infty} \frac{(-1)^n}{\frac{1}{p} - n} + p + \sum_{n=1}^{\infty} \frac{(-1)^n}{\frac{1}{p} + n} = \frac{\pi}{\sin \frac{\pi}{p}}.
\end{aligned}$$

To prove the last equality one can show that

$$\varphi(z) = \frac{\pi}{\sin \pi z} - \frac{1}{z} - \sum_{n=1}^{\infty} (-1)^n \left(\frac{1}{z-n} + \frac{1}{z+n} \right)$$

is an entire and bounded holomorphic function (we put $\varphi(z) = 0$ when z is an integer) and therefore (by Liouville's theorem) it is identically zero. For $z = 1/p$ we obtain the desired equality. There is also another way to prove Lemma 3—the integral $\int_0^{\infty} dt/t^{1/p}(1+t)$ can be done by contour integration (see chapter 3.6, the fourth type of integrals, H. Cartan *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Hermann, Paris, 1961). \square

Therefore, using similar reasoning we can extend Proposition 1 as follows.

Proposition 2. For $p, q > 1$ such that $1/p + 1/q = 1$ and sequences of non-negative numbers $(a_m), (b_n)$ such that $\sum_{m=1}^{\infty} a_m^p, \sum_{n=1}^{\infty} b_n^q$ are convergent

$$\sum_{m,n=1}^{\infty} \frac{a_m b_n}{m+n} \leq \frac{\pi}{\sin \frac{\pi}{p}} \left(\sum_{m=1}^{\infty} a_m^p \right)^{1/p} \left(\sum_{n=1}^{\infty} b_n^q \right)^{1/q}. \quad (2)$$

The proof is essentially as above (we use Hölder's inequality instead of Schwarz's inequality). The constant in (2) is also optimal.

This note grew out of useful discussion with Paweł Strzelecki, to whom I am greatly indebted. I would like to thank Jan Herczyński and Michał Wojciechowski for their help in preparing this text.

*Instytut Matematyki
Uniwersytet Warszawski
ul. Banacha 2
02-097 Warszawa, Poland*

A Note on Fubini's Theorem

Camille Debieve

The purpose of this note is to give a very simple example of a nonnegative function of two real variables $f(x, y)$ such that

$$\int_0^1 \left(\int_0^1 f(x, y) dx \right) dy = \int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = 0$$

and such that

$$\iint_{[0,1] \times [0,1]} f dx dy$$

does not exist in the Riemann sense.

Gelbaum and Olmsted give two examples of such functions in “Counterexamples in Analysis (Holden-Day Inc.)”.

The first one is the characteristic function of a subset A of the unit square $[0, 1] \times [0, 1]$ that is dense in the unit square and such that every vertical or horizontal line meets A in only one point. The construction of such a set is based on a construction by stages.

The second one uses a non-measurable set of the plane having at most two points in common with a line. The construction of such a set is essentially based on Zorn's lemma.

These two examples are rather hard to understand by undergraduate students.

We can meet the same requirements with the following very simple example:

For $k = 1, 2, 3, \dots$, let A_k be the set of all points $(m/2^k, n/2^k)$ in the unit square, where m and n are odd integers, and put $A = \bigcup_{k=1}^{\infty} A_k$.

The finite sets A_k have disjoint projections into the coordinate axes. Therefore A has at most finitely many points on each vertical or horizontal line. Letting f be the characteristic function of A , it follows that

$$\int_0^1 f(x, y) dx = 0 = \int_0^1 f(x, y) dy.$$

On the other hand, A is dense in the unit square, so that the upper integral of f is equal to one and its lower integral is zero which shows that f is not integrable in the Riemann sense.

*Université Catholique de Louvain
Département de Mathématiques
2, Chemin du Cyclotron
B 1348 Louvain-la-Neuve
Belgium.
E-mail: debieve@amm.ucl.ac.be*

UNSOLVED PROBLEMS

Edited by: **Richard Guy**

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

Parker's Permutation Problem Involves the Catalan Numbers

The late E. T. Parker submitted the following problem to the MONTHLY:

Let a_i , $1 \leq i \leq n$ be n integers whose sum is a multiple of $n + 1$. Must there be permutations $\{x_i\}$ and $\{y_i\}$ of $1, 2, \dots, n$ such that, for each i ,

$$x_i + y_i \equiv a_i \pmod{n + 1}?$$

Neither the proposer nor the referees gave a solution, so the problem was not used in the regular Problems section.

John Selfridge recasts the problem in the following form:

Write out the addition table for the numbers 1 up to n , mod $n + 1$; e.g., for $n = 5$:

	5	4	3	2	1
1	0	5	4	3	2
2	1	0	5	4	3
3	2	1	0	5	4
4	3	2	1	0	5
5	4	3	2	1	0

where there are n entries 0 and $n - 1$ of each of the others. Given any n of these residues, with repetitions allowed, but with zero sum mod $n + 1$, can they be chosen just one from each row and one from each column?

It is not difficult to verify that this can always be done if $n \leq 5$ and a computer could probably go twice as far. But after looking at the problem for a while, you'll feel sure that the answer is "yes"—but can you prove it?

Clearly there are $n!$ choices from the table, but these include duplicates. Just how many multisets of n residues are there, mod $n + 1$?

It is easy to make the rough estimate that strongly suggests that the answer to Parker's problem is affirmative. Precisely the number is

$$\sum_{k=0}^{n-1} p(k(n+1); n, n)$$

where $p(m; n, n)$ is the number of partitions of m into at most n parts each no bigger than n . Here is a table of the summand for $1 \leq n \leq 7$.

n											Total
1											1
2											2
3											5
4											14
5											42
6											132
7											429

Reminiscent of the Stirling numbers of the second kind, except that the present numbers are a bit smaller, and the row sums, instead of being Bell numbers, are the Catalan numbers,

$$\frac{1}{n+1} \binom{2n}{n}.$$

To prove that this is so, Ira Gessel recalls that the number of partitions of k with at most m parts, each no bigger than n , is the coefficient of q^k in the q -binomial coefficient

$$\left[\begin{matrix} m+n \\ n \end{matrix} \right] = \frac{(q)_{m+n}}{(q)_m (q)_n}$$

where $(q)_i = (1-q)(1-q^2) \cdots (1-q^i)$. What we want is the sum of the coefficients of q^k , for k divisible by $n+1$, in $\left[\begin{matrix} 2n \\ n \end{matrix} \right]$.

Suppose that $P(q)$ is any polynomial and let r be a positive integer. To find the sum of the coefficients of q^k for k divisible by r in $P(q)$, let ζ be a primitive r th root of unity. Then the desired sum is easily seen to be

$$\frac{1}{r} (P(1) + P(\zeta) + P(\zeta^2) + \cdots + P(\zeta^{r-1})).$$

(A similar formula can be used to find the sum of the coefficients of q^k in $P(q)$ for k in a given residue class modulo r .)

Return to our original problem and let ζ be a primitive $(n+1)$ st root of unity and let $P(q) = \left[\begin{matrix} 2n \\ n \end{matrix} \right]$. It is sufficient to show that for $i = 1, \dots, n$, $P(\zeta^i) = 0$. But ζ^i is some primitive d th root of unity for some divisor d of $n+1$. So it is enough to show that if d is a divisor of $n+1$ with $d > 1$ and ξ is a primitive d th root of unity, then $P(\xi) = 0$. Let $\phi_d(q)$ be the d th cyclotomic polynomial. It is sufficient to show that $P(q)$ is divisible by $\phi_d(q)$. Equivalently we may show that the power of $\phi_d(q)$ dividing $(q)_{2n}$ is strictly greater than the power of $\phi_d(q)$ dividing $(q)_n^2$. The power of $\phi_d(q)$ dividing $(q)_i$ is $\lfloor i/d \rfloor$ so we need only show that

$$\left\lfloor \frac{2n}{d} \right\rfloor > 2 \left\lfloor \frac{n}{d} \right\rfloor.$$

But, if $n+1 = td$ with $d > 1$, then the left side is $2t-1$ and the right side is $2t-2$.

Gessel also observes that the same argument shows that the sum of the coefficients of $\left[\begin{smallmatrix} 2n \\ n \end{smallmatrix} \right]$ congruent to j modulo $n + 1$ is also the Catalan number for any j . In some cases $n + 1$ can be replaced by other numbers, in the sense that if s is the number replacing $n + 1$, the sum of the coefficients of $\left[\begin{smallmatrix} 2n \\ n \end{smallmatrix} \right]$ congruent to j modulo s is independent of j , and is thus equal to $\frac{1}{s} \binom{2n}{n}$, but he hasn't investigated this further.

Apart from the Catalan numbers, none of the sequences associated with our triangular array appear to be in the second edition of Sloane's Handbook [3], unless a preprint of this paper reaches him in time and some of them receive his seal of approval. They can be written in terms of the partition function, but as one leaves the edges of the table, the formulas become less and less closed. For $k = 1$ and $k = n - 2$ they are $p(n + 1) - 2$ and $p(n + 2) - 4$, where $p(m)$ is the number of partitions of m . If we write $P(m) = \sum_{i=1}^m p(i)$, then the formulas for $k = 2$ and $k = n - 3$ are $p(2n + 2) - 2P(n + 1) + 1$ and $p(2n + 3) - 2P(n + 2) + 7$. Have readers seen this array in any other context?

Let us return to Parker's original problem. For $n = 1$ and 2 there are just the right number of choices. For $n = 3$ the choice 013 can be made in two ways. For $n = 4$ the choices

0014, 0023, 1234 can each be made in 3 ways;
0113, 0122, 0244, 0334 each in 2 ways; and
0000, 1112, 1144, 1333, 2224, 2233, 3444 uniquely.

For $n = 5$, 01245 can be chosen in 8 ways;
each of 00123, 00345, 01344, 02235 in 6 ways;
00015, 00024, 01335, 02334, 11235, 12234, 13455, 23445 in 4 ways;
00114, 00255, 01122, 01155, 02244, 04455, 11334, 23355 in 3 ways;
eleven others in 2 ways and the remaining ten uniquely.

Clearly the number of repetitions is larger the more distinct the sizes of the parts. Is there any neat way of relating the number of repetitions to the shape of the partition?

There are so many manifestations of the Catalan numbers [1, 2] that it seems likely that there are more direct combinatorial proofs awaiting discovery. And what about Parker's original problem? There's a slightly unaesthetic feature, which we tried to eliminate by rewording it, but some clash between n and $n + 1$ remains. Jerrold Griggs avoids it by extending the number of numbers to $n + 1$, so that $\sum_{i=1}^{n+1} a_i$ is a multiple of $n + 1$; Parker's problem is the case $a_{n+1} = 0$, but Griggs believes that the more general conjecture is also true.

REFERENCES

1. Henry W. Gould, *Bell & Catalan Numbers: research bibliography of two special number sequences*, 6th edition, Morgantown WV, 1985.
2. Michael J. Kuchinski, *Catalan Structures and Correspondences*, MSc thesis, West Virginia University, 1977.
3. Neil J. A. Sloane, *The New Book of Integer Sequences*, W. H. Freeman, 1993.

PROBLEMS AND SOLUTIONS

Edited by:

Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before August 31, 1993 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10290. *Proposed by David Allison, University of Cape Town, Rondebosch, South Africa.*

Let $c \in \mathbb{N}$. Consider the expression $S_c(n) = \sum_{r=1}^n r^c$.

- (a) Show that $S_c(n)/S_1^2(n)$ is a polynomial in $S_1(n)$ when c is odd and $c > 1$.
- (b) Show that $S_c(n)/S_2(n)$ is a polynomial in $S_1(n)$ when c is even.

10291. *Proposed by Howard Morris, Chatsworth, CA.*

Let k be a positive integer and let $\langle x_n \rangle$ be a nondecreasing sequence of real numbers with $x_1 > 1$ for which $\sum (1/x_n)$ converges. Show that $\sum (\ln x_n)^k / x_n$ converges if and only if $\sum (\ln n)^k / x_n$ converges.

10292. *Proposed by Jean Anglesio, Garches, France.*

Obtain explicit values for the following series.

$$(a) \sum_{n=1}^{\infty} \arctan\left(\frac{2}{n^2}\right)$$

$$(b) \sum_{n=1}^{\infty} \arctan\left(\frac{8n}{n^4 - 2n^2 + 5}\right)$$

10293. *Proposed by Moshe Rosenfeld, Pacific Lutheran University, Tacoma, WA.*

Suppose four distinct lines through the origin in \mathbb{R}^3 have the property that the six acute angles between pairs of these lines are all equal. Prove that this configuration of four lines is isometric either to the diagonals of a cube or to a configuration of four of the six diagonals of a regular icosahedron.

10294. *Proposed by Derek A. Holton, University of Otago, Dunedin, New Zealand.*

Given a positive integer n , let $N = \{1, \dots, n\}$. For a positive integer k , say that n is k -good if N can be partitioned into k sets each with the same sum. Show that n is k -good if k divides $\binom{n}{2}$ and n is sufficiently large.

10295. *Proposed by Mark A. Pinsky, Northwestern University, Evanston, IL.*

Let $f: \mathbb{R}^3 \rightarrow \mathbb{R}^1$ be continuous in the ball $\{x : |x| \leq a\}$ with continuous first partial derivatives. The Fourier transform, denoted $\hat{f}(\mu)$, is given by the expression

$$(2\pi)^{-3} \int_{\{x: |x| \leq a\}} f(x) e^{-i\langle \mu, x \rangle} dx.$$

Prove that

$$f(0) = \lim_{R \uparrow \infty} \int_{\{\mu: \mu \leq R\}} \hat{f}(\mu) d\mu$$

if and only if the surface integral $\int_{\{x: |x|=a\}} f d\sigma = 0$.

10296. *Proposed by Paul Erdős, Hungarian Academy of Sciences, Budapest, Hungary.*

Let $f(n, a)$ denote the least common multiple of the n consecutive integers between a and $a + n - 1$ inclusive.

(a) Show that for any positive integer k , there is a number $n_0(k)$ such that for each integer $n > n_0(k)$, there is an integer a with

$$f(n, a) > f(n, a + 1) > \dots > f(n, a + k).$$

(b)* Do there exist infinitely many integers n for which there is a pair of integers a, b with $a < b$ and $f(n, a) > f(n + 1, b)$?

10297. *Proposed by Zalman Rubinstein, University of Haifa, Haifa, Israel.*

Let $p(z)$ be a polynomial of degree n .

(a) Show that $p(z)$ can be written as a sum of four polynomials $q_0(z), q_1(z), q_2(z), q_3(z)$, each of degree at most n with all roots of all $q_i(z)$ lying on the unit circle $\{z: |z| = 1\}$.

(b)* Is there a polynomial $p(z)$ which cannot be expressed as a sum of fewer than 4 such $q_i(z)$?

Notes: (10294) This problem was suggested by the treatment of the case $k = 2$ in Harris Schulz, “Summation properties of $\{1, \dots, n\}$ ”, *Mathematical Spectrum*, 23 (1990/1), 8–11. **(10296)** Explicit estimates on $n_0(k)$ are desired. **(10297)** A polynomial all of whose roots lie on the unit circle is known as a C -polynomial.

SOLUTIONS

Flattening an Integer Sequence

E3267 [1988, 456]. *Proposed by Barry Hayes, Donald Knuth, and Carlos Subi, Stanford University, CA.*

Given a sequence (x_1, x_2, \dots, x_l) of nonnegative integers in which $x_k > 1$ for some k , where $1 < k < l$, let us say that a “ k -move” is the operation of replacing the subsequence (x_{k-1}, x_k, x_{k+1}) by $(x_{k-1} + 1, x_k - 2, x_{k+1} + 1)$.

(a) Prove that repeated application of such moves to the sequence $(0^m, 2m, 0^m)$ always leads to the sequence $(1^m, 0, 1^m)$ after exactly $\frac{1}{3}(m+1)(m+\frac{1}{2})m$ moves. Here 0^m and 1^m stand for sequences of m 0’s and m 1’s, respectively.

(b) Prove that, for sufficiently large m , the starting sequence $(0^m, a_1, \dots, a_n, 0^m)$ leads inexorably to the sequence $(0^{m+p}, 1^q, 0, 1^r, 0^{m+n-p-q-r-1})$ for some p, q , and r , if a_1, \dots, a_n are positive integers. Furthermore, p, q , and r can be expressed in terms of $\sum_{j=1}^n a_j$ and $\sum_{j=1}^n ja_j$. However, how many moves does this transformation require?

Solution by Albert Nijenhuis, Seattle, WA. We use the term *state* to mean a (doubly) infinite sequence $X = \{x_i\}_{i=-\infty}^{\infty}$ such that each x_i is a non-negative integer and $\sum |x_i| < \infty$. The finite sequences occurring in the problem can be thought of as abbreviations for the states obtained from them by appending a string of zeros on each end. If X is a state, the *support* of X , denoted by $\text{supp}(X)$, is the minimal interval $[a, b]$ such that $x_i = 0$ for all i not in $[a, b]$.

If $X = \{x_i\}_{i=-\infty}^{\infty}$ is a state and if $x_k > 1$ for a particular value of k , the k -move σ_k can be defined by requiring the state $\sigma_k(X)$ to be the same as X except that the entries x_{k-1}, x_k, x_{k+1} are replaced by $x_{k-1} + 1, x_k - 2, x_{k+1} + 1$ respectively. If $x_k \leq 1$, of course $\sigma_k(X)$ is undefined.

A state X is *terminal* if $\sigma_k(X)$ is undefined for all integers k , i.e., if $x_i \in \{0, 1\}$ for all i . A state X is *coherent* if its support does *not* contain a subinterval $[a, b]$ such that $a < b$, $x_a = x_b = 0$, and $x_i = 1$ whenever $a < i < b$. The states occurring in parts (a) and (b) of the problem are coherent.

We shall find it useful to consider the moments $m_j(X) = \sum_{i=-\infty}^{\infty} i^j x_i$ for $j = 0, 1, 2$. If $\sigma_k(X)$ is well-defined, clearly $m_j(\sigma_k(X)) - m_j(X) = (k-1)^j - 2k^j + (k+1)^j$.

The following ten assertions are sufficiently easy to prove that we shall omit all proofs except for that of assertion (iv) in the interests of brevity.

- (i) If X is a state and if $\sigma_k(X)$ is well-defined, then $\text{supp}(X) \subset \text{supp}(\sigma_k(X))$.
- (ii) If X is a state and if $\sigma_k(X)$ is well-defined, then $m_0(\sigma_k(X)) = m_0(X)$ and $m_1(\sigma_k(X)) = m_1(X)$. Thus k -moves do not change the center of gravity $\mu(X) = m_1(X)/m_0(X)$.

- (iii) If X is a state and if $\sigma_k(X)$ is well-defined, then $m_2(\sigma_k(X)) = m_2(X) + 2$.
 (iv) If X is a coherent state, then $m_2(X)$ is bounded above by a function of $m_0(X)$ and $m_1(X)$.

Proof: Let $[a, b] = \text{supp}(X)$. Define $\phi(k) = \sum_{i=a}^k x_i - (k - a)$ for $k \in [a, b]$. Clearly $\phi(k) \geq \phi(k - 1) - 1$ with equality only if $x_k = 0$. We claim that $\phi(k) \geq 0$ for all k in $[a, b]$. Indeed suppose k is the first integer in $[a, b]$ for which $\phi(k) < 0$. Then $k \geq a + 2$, $x_k = 0$, $\phi(k) = -1$, and $\phi(k - 1) = 0$. Let j be the last integer in $[a, k]$ for which $\phi(j) > 0$. Then $j < k - 1$, $\phi(j + 1) = 0$ and $\phi(j) = 1$, so that $x_{j+1} = 0$ and $x_i = 1$ for $j + 1 < i \leq k - 1$. (If $j + k = k - 1$, there are no such values of i .) Thus $[j + 1, k]$ violates the coherence of X . Hence $\phi(k) \geq 0$ for all k in $[a, b]$ and in particular $m_0(X) - (b - a) = \phi(b) \geq 0$. Thus $b - a \leq m_0(X)$. Hence $|i| \leq |m_1(X)/m_0(X)| + m_0(X)$ for $i \in \text{supp}(X)$ and we have the crude bound

$$m_2(X) \leq (|m_1(X)/m_0(X)| + m_0(X))^2 m_0(X).$$

- (v) If X is a coherent state and if $\sigma_k(X)$ is well-defined, then $\sigma_k(X)$ is coherent.

- (vi) Given a coherent state X , all sequences of well-defined states

$$X, \sigma_{k_1}(X), \sigma_{k_2}(\sigma_{k_1}(X)), \dots$$

are bounded in length; every such sequence of maximal length ends with a terminal state.

- (vii) A non-zero coherent terminal state $T = \{t_i\}_{i=-\infty}^{\infty}$ is of one of two types:

Type 1. $\text{supp}(T) = [a, b]$ and $t_i = 1$ for all $i \in [a, b]$. In this case we have

$$m_0(T) = b - a + 1 \quad m_1(T) = (a + b)(b - a + 1)/2$$

$$\mu(T) = m_1(T)/m_0(T) = (a + b)/2$$

$$m_2(T) = b(b + 1)(2b + 1)/6 - (a - 1)a(2a - 1)/6.$$

Type 2. $\text{supp}(T) = [a, b]$ and there is an integer c , $a < c < b$, such that $t_c = 0$ and $t_i = 1$ for $i \neq c$, $i \in [a, b]$. In this case we have

$$m_0(T) = b - a \quad m_1(T) = (a + b)(b - a + 1)/2 - c$$

$$\mu(T) = m_1(T)/m_0(T) = (a + b)/2 - (c - (a + b)/2)/(b - a)$$

$$m_2(T) = b(b + 1)(2b + 1)/6 - (a - 1)a(2a - 1)/6 - c^2.$$

(viii) If T is a non-zero coherent terminal state with $\text{supp } T = [a, b]$, then $\mu(T)$ is at distance less than $1/2$ from $(a + b)/2$ and is a rational number expressible with denominator $m_0(T)$. If T is of type 1, then either $m_0(T)$ is odd and $\mu(T)$ is an integer or $m_0(T)$ is even and $\mu(T)$ is half an odd integer. If T is of type 2 and $\mu(T)$ is an integer, then $\mu(T) = (a + b)/2$, $m_0(T)$ is even, and the “zero” element of T lies exactly in the middle. When T is of type 2, $\mu(T)$ is never half an odd integer.

(ix) Non-zero coherent terminal states T are uniquely determined by the values of $m_0(T)$ and $m_1(T)$.

(x) The number of steps from a coherent state X to its unique terminal state T via k -moves is $(m_2(T) - m_2(X))/2$.

Part (a) of the problem now follows. Here $X = \{x_i\}_{i=-\infty}^{\infty}$ is given by $x_i = 0$ for $i \neq 0$ and $x_0 = 2m$, where m is a given positive integer. Thus $m_0(X)$ is even,

$\mu(X) = 0$, and X is coherent. Hence for the resulting terminal state $T = \{t_i\}_{i=-\infty}^{\infty}$ we have $\mu(T) = 0$ and $m_0(T)$ even. Thus by (viii) T is of type 2, $t_i = 1$ for $1 \leq |i| \leq m$, and $t_i = 0$ otherwise. The number of moves needed to go from X to T is then $(m_2(T) - m_2(X))/2 = 1^2 + 2^2 + \cdots + m^2 = (m+1)(m+1/2)m/3$, so that Part (a) is established.

Now we turn to Part (b). Here $x_i = a_i$ for $1 \leq i \leq n$ and $x_i = 0$ otherwise. For convenience we put $m_j = m_j(X) = \sum_{i=1}^n i^j a_i$. If T is the corresponding terminal state, then $m_0 = m_0(T)$, $m_1 = m_1(T)$, but $m_2 \neq m_2(T)$ unless $a_i = 1$ for $i = 1, 2, \dots, n$. The characterization of the terminal state T and the determination of the quantities p, q, r , and N will now be sketched.

If m_0 is odd and m_1/m_0 is an integer, or if m_0 is even and m_1/m_0 is half an odd integer, then by (viii) we have a terminal state of type 1. If $\text{supp}(T) = [a, b]$, then from (vii) and (viii) we have $(a+b)/2 = m_1/m_0$ and $(b-a)/2 = (m_0-1)/2$. Since $p = a-1$ and $q = b-a+1$, the quantities p, q, r of the problem are given by

$$p = m_1/m_0 - (m_0 + 1)/2 \quad q = m_0 \quad r = 0$$

and if N is the number of moves needed to go from X to T , then

$$2N = m_1^2/m_0 + (m_0^3 - m_0)/12 - m_2.$$

In all other cases (viii) tells us that the terminal state T is of type 2. If $\text{supp}(T) = [a, b]$, then from (vii) and (viii) we have $(b-a)/2 = m_0/2$ and $(a+b)/2 = m_1/m_0 + 1/2 - f$, where $0 < f < 1$. Since

$$a = (a+b)/2 - (b-a)/2 = m_1/m_0 - (m_0-1)/2 - f$$

and a is an integer, f must be the fractional part of $m_1/m_0 - (m_0-1)/2$.

Routine calculation now gives

$$p = m_1/m_0 - (m_0 + 1)/2 - f, \quad q = m_0(1-f), \quad r = m_0 f,$$

and simplification of $2N = m_2(T) - m_2(X)$ yields

$$2N = m_1^2/m_0 + (m_0^3 - m_0)/12 - m_2 + f(1-f)(m_0^2 + m_0).$$

When $f = 0$, these formulas reduce to those obtained for type 1. Also, when $n = 1$ and $a_1 = 2m$, these formulas reduce to those of part (a).

Solved also by A. M. Adelberg, S. F. Barger, G. W. Peck, P. S. Zwier, and the proposers.

Minuscule Sets of Real Numbers

6634 [1990, 535]. *Proposed by R. W. Zeamer, University College, London, England.*

Let S be the set of real numbers of the form

$$\sum_{n=2}^{\infty} \varepsilon_n/n!,$$

where $\varepsilon_n = 0$ or 1. Let K be the field generated by S . Is it true that $K = \mathbb{R}$?

Solution by O. P. Lossers, Eindhoven University of Technology, The Netherlands.
It is not true that $K = \mathbb{R}$. To prove this we introduce the notion of a *minuscule* set of real numbers. We shall prove that K is minuscule but \mathbb{R} is not.

A set V of real numbers will be called *elementary minuscule* iff for every $\delta > 0$ there exists a positive number N with the following property: given $M \geq N$ there are compact intervals V_1, \dots, V_μ such that $\mu = \lfloor M^\delta \rfloor$, $V \subseteq \bigcup_{i=1}^\mu V_i$, and $|V_i| \leq 1/M$ for each i , where $|V_i|$ stands for the length of V_i . A set V of real numbers will be called *minuscule* iff it is contained in a countable union of elementary minuscule sets. We shall use the following five properties of minuscule sets.

(i) If V is minuscule, then V is meager, i.e., is expressible as the union of countably many nowhere dense sets; in particular, we can conclude that $V \neq \mathbb{R}$.

(ii) If V is minuscule and $W \subset V$, then W is minuscule.

(iii) If V_1, V_2, \dots are minuscule, then $\bigcup_{i=1}^\infty V_i$ is minuscule.

(iv) If U is open, if $f: U \rightarrow \mathbb{R}$ is continuously differentiable, if $V \subseteq U$, and if V is minuscule, then $f(V)$ is minuscule.

(v) If V and W are minuscule, then $V + W = \{v + w: v \in V, w \in W\}$ is minuscule.

Assertions (i), (ii), and (iii) are immediate. To prove (iv) we may assume that V is elementary minuscule. Let C be a compact interval contained in U . On C the function f' is bounded, say $|f'| \leq A$. Suppose $V \subseteq \bigcup_{i=1}^\mu V_i$, where $\mu = \lfloor M^\delta \rfloor$ and $|V_i| \leq 1/M$ for $i = 1, 2, \dots, \mu$. Then $f(C \cap V) \subseteq \bigcup_{i=1}^\mu f(C \cap V_i)$, where $f(C \cap V_i)$ is a compact interval of length at most A/M . Since $M^\delta \leq (M/A)^{2\delta}$ if M is large enough, this proves that $f(C \cap V)$ is elementary minuscule. As U can be written as a countable union of compact intervals, we can conclude that $f(V)$ is minuscule.

To prove (v) we may assume that V and W are elementary minuscule. So suppose $V \subseteq \bigcup_{i=1}^\mu V_i$, where $\mu = \lfloor M^\delta \rfloor$, $|V_i| \leq 1/M$ for $i = 1, 2, \dots, \mu$, $W \subseteq \bigcup_{j=1}^\mu W_j$, and $|W_j| \leq 1/M$ for $j = 1, 2, \dots, \mu$. Then $V_i + W_j$ is an interval with $|V_i + W_j| \leq 2/M$. The number of such intervals $V_i + W_j$ is $\mu^2 = \lfloor M^\delta \rfloor^2$, which is less than $(M/2)^{4\delta}$ if M is large enough. This proves that $V + W$ is minuscule.

Using the assertions (i)–(v) we proceed to prove that K is minuscule. The collection of compact intervals

$$\left\{ \left[\sum_{n=2}^s \frac{\varepsilon_n}{n!}, \frac{e}{s!} + \sum_{n=2}^s \frac{\varepsilon_n}{n!} : \varepsilon_2, \dots, \varepsilon_s \in \{0, 1\} \right] \right\}$$

is a covering of S with 2^{s-1} intervals of length $e/s!$. Since $2^{s-1} < (s!/e)^\delta$ if s is large enough, we conclude that S is elementary minuscule.

Now we define an infinite sequence of minuscule sets $\{S_n\}_{n=0}^\infty$ by putting $S_0 = S$ and

$$\begin{aligned} S_{n+1} = & S_n \cup \{-a: a \in S_n\} \cup \{1/a: a \in S_n, a \neq 0\} \\ & \cup \{a + b: a \in S_n, b \in S_n\} \cup \{a^2/2: a \in S_n\}. \end{aligned}$$

If $H = \bigcup_{i=0}^\infty S_i$, then H is minuscule from properties (i)–(v). Also H is a field; for instance from $a, b \in H$ follows $ab = (a + b)^2/2 - a^2/2 - b^2/2 \in H$. Thus $H = K$. Hence K is minuscule. By (i) \mathbb{R} is not minuscule and so $K \neq \mathbb{R}$.

We remark that in place of (i) we could use the assertion that a minuscule set has Lebesgue measure zero.

Solved also by Shaw Chen, Howard Morris, and M. J. Pelling. The solution by Pelling accompanied the original proposal.

The Average Size of a Certain Arithmetic Function

6660 [1991, 446]. Proposed by Richard Alan Gillman, Valparaíso University, Valparaíso, IN.

Let f be the multiplicative arithmetic function defined by $f(1) = 1$ and $f(p^a) = pf(a)$ for all primes p and all positive integers a .

- (i) Prove that $0 < f(n) \leq n$ for all n .
- (ii) Prove that the ratio

$$\frac{\sum_{j=1}^n f(j)}{\sum_{j=1}^n j}$$

has a positive limit C as $n \rightarrow \infty$.

- (iii) Estimate the difference

$$\sum_{j=1}^n f(j) - C \sum_{j=1}^n j.$$

Solution by Rainer Tschiersch, Johann Wolfgang Goethe-Universität, Frankfurt am Main, Germany. We first establish (i) by induction on n . Suppose that n is a positive integer greater than 1 and that the assertion of (i) is true for all integers k with $1 \leq k < n$. If n is a power of some prime, say $n = p^a$, then, since $a < n$, it follows from the inductive hypothesis that $0 < f(n) = f(p^a) = pf(a) \leq pa \leq p^a = n$. Otherwise there exist relatively prime integers n_1, n_2 with $1 < n_1, n_2 < n$ and $n = n_1 n_2$; but in this case, because of the multiplicativity of f , it follows that $0 < f(n) = f(n_1)f(n_2) \leq n_1 n_2 = n$. Thus (i) is proved.

Define the arithmetical function h by $f = g * h$, where $*$ denotes Dirichlet convolution and g is the identity function, i.e., $g(n) = n$ for all positive integers n . We prove (ii) with $C = \sum_{n=1}^{\infty} h(n)/n^2 = 0.835107636 \dots$; for (iii) we provide the estimate $O(n^{3/2} \log n)$ with a specific O -constant.

Since g is completely multiplicative, its inverse under convolution is $\mu \cdot g$, where μ is the Möbius function and the dot indicates pointwise multiplication. Hence h is given by $h = f * (\mu \cdot g)$. Thus h is multiplicative and its values at prime-powers are

$$h(p) = f(p) - p = 0,$$

$$h(p^b) = f(p^b) - pf(p^{b-1}) = pf(b) - p^2 f(b-1) \quad (b \geq 2).$$

Using (i) we obtain the rough estimates

$$h(p^b) \leq f(p^b) \leq p^b,$$

$$h(p^b) \geq -pf(p^{b-1}) \geq -p \cdot p^{b-1} = -p^b.$$

Hence $|h(n)| \leq n$, and $h(n) = 0$ if n is not square-full. (A positive integer n is called square-full if no prime appears in its canonical factorization with multiplicity 1, i.e., if $n = 1$ or if $p^2 | n$ for each prime factor p of n .)

Since each square-full number is uniquely expressible as $r^2 s^3$ where s is square-free, we have

$$\sum_{n=1}^{\infty} |h(n)| n^{-1-\sigma} \leq \sum_{n \text{ square-full}} n^{-\sigma} = \zeta(2\sigma)\zeta(3\sigma)/\zeta(6\sigma) < \infty$$

for any real $\sigma > 1/2$, where ζ denotes the Riemann zeta function. From now on let σ be a number in $(\frac{1}{2}, 1)$.

If y is any positive real number, we have

$$(y-1)y/2 < \sum_{m \leq y} m = [y]([y] + 1)/2 \leq y(y+1)/2,$$

so that

$$\sum_{m \leq y} g(m) = \sum_{m \leq y} m = y^2/2 + R(y), \quad |R(y)| \leq y/2.$$

Thus the convolution relation $f = g * h$ gives

$$\begin{aligned} \sum_{j \leq x} f(j) &= \sum_{d \leq x} h(d) \sum_{m \leq x/d} g(m) \\ &= (x^2/2) \cdot \sum_{d \leq x} h(d) d^{-2} + \sum_{d \leq x} h(d) \cdot R(x/d) \\ &= (x^2/2) \cdot \sum_{d=1}^{\infty} h(d) d^{-2} - (x^2/2) \sum_{d > x} h(d) d^{-2} + \sum_{d \leq x} h(d) \cdot R(x/d) \\ &= (x^2/2) \cdot C + R_1(x) + R_2(x). \end{aligned}$$

Here

$$|R_1(x)| \leq (x^2/2) \cdot \sum_{d > x} d^{\sigma-1} |h(d)| d^{-1-\sigma} \leq (x^{1+\sigma}/2) \cdot \sum_{d > x} |h(d)| d^{-1-\sigma}$$

and, since $|R(y)| \leq y/2 \leq y^{1+\sigma}/2$ for $y \geq 1$,

$$|R_2(x)| \leq \sum_{d \leq x} |h(d)| (1/2)(x/d)^{1+\sigma} = (x^{1+\sigma}/2) \cdot \sum_{d \leq x} |h(d)| d^{-1-\sigma}.$$

Combining these two estimates, we have

$$|R_1(x)| + |R_2(x)| \leq (x^{1+\sigma}/2) \cdot \sum_{d=1}^{\infty} |h(d)| d^{-1-\sigma}.$$

On the other hand, since $C = \sum h(d) d^{-2} \leq \sum |h(d)| d^{-1-\sigma}$, we have (for $x \geq 1$)

$$C|R(x)| \leq (x/2) \cdot \sum_{d=1}^{\infty} |h(d)| d^{-1-\sigma} \leq (x^{1+\sigma}/2) \cdot \sum_{d=1}^{\infty} |h(d)| d^{-1-\sigma}.$$

Thus

$$\begin{aligned} \left| \sum_{j \leq x} f(j) - C \sum_{j \leq x} j \right| &= |R_1(x) + R_2(x) - CR(x)| \\ &\leq x^{1+\sigma} \sum_{d=1}^{\infty} |h(d)| d^{-1-\sigma} \\ &\leq x^{1+\sigma} \zeta(2\sigma) \zeta(3\sigma) / \zeta(6\sigma) \end{aligned} \tag{1}$$

for any σ in $(\frac{1}{2}, 1)$.

Suppose now that $x \geq 8$ and let us take $\sigma = 1/2 + 1/\log x$ in (1), a choice which essentially minimizes the right-hand side. Using the inequalities

$$\zeta(2\sigma) < 2\sigma/(2\sigma - 1) < 2/(2\sigma - 1)$$

and

$$\zeta(3\sigma)/\zeta(6\sigma) \leq \zeta(3/2)/\zeta(3),$$

we obtain (for $x \geq 8$)

$$\left| \sum_{j \leq x} f(j) - C \sum_{j \leq x} j \right| < (e\zeta(3/2)/\zeta(3))x^{3/2} \log x \\ < 6x^{3/2} \log x. \quad (2)$$

The inequality (2) is a bound of the sort solicited in (iii) and provides the following more specific version of (ii)

$$\sum_{j \leq x} f(j) \Big/ \sum_{j \leq x} j - C = O(x^{-1/2} \log x) \quad (x \rightarrow \infty).$$

Editorial comment. The editors wish to thank Kevin Ford for calculating the numerical value of $C = 0.835107636\dots$ given above. The conclusions obtained above assert that on average $f(n)/n$ is around 0.8351, even though the ratios $f(n)/n$ appear to be everywhere dense in $(0, 1)$.

The order of magnitude of the estimate in (2) can be lowered slightly by referring to known results on the distribution of square-free numbers. Instead of writing $f = g * h$, where g is the identity function as above, we may write $f = G * H$, where $G(n) = n|\mu(n)|$, i.e., where $G(n) = n$ if n is square-free and $G(n) = 0$ otherwise. The inverse of G under convolution is easily seen to be the completely multiplicative arithmetic function R such that $R(p) = -p$ for all primes p . Hence $H = f * R$. Thus H is multiplicative and its values at prime powers are given by

$$H(p^k) = \sum_{j=0}^k f(p^{k-j})(-p)^j = \sum_{j=0}^{k-1} f(k-j)(-1)^j p^{j+1} + (-p)^k$$

for p a prime and k a positive integer. Thus

$$H(p) = 0, \quad H(p^k) = \sum_{j=0}^{k-2} (-1)^j f(k-j)p^{j+1} \quad (k \geq 2).$$

It is not hard to show that $\sum |H(n)|n^{-1-\varepsilon}$ converges for any positive ε , and it is known that for x large

$$\sum_{n \leq x} G(n) = 3\pi^{-2}x^2 + o(x^{3/2}). \quad (3)$$

Using these two assertions and proceeding as in the above solution, we easily find

$$\sum_{n \leq x} f(n) = (C/2)x^2 + o(x^{3/2}). \quad (4)$$

More specific results on the error term in (3) can be used to get a more specific error term in (4). In particular, Vinogradov's zero-free region for $\zeta(2s)$ can be used to improve the error term in (3) to $O(x^{3/2} \exp(-c(\log x)^{2/3}(\log \log x)^{-1/3}))$.

The proposer gave a solution of (i) and the editors provided solutions of (ii) and (iii). The solution given above was the only other one received.

A Guessing Game

E 3448 [1991, 553]. *Proposed by Howard Taylor, University of Delaware, Newark.*

Each day the call-in program on a local radio station conducts the following game. A number is drawn at random from $\{1, 2, \dots, n\}$. Callers are supposed to guess the number drawn. If a caller guesses the correct number, the station awards

a prize and the game is finished. If a caller guesses incorrectly, the station announces the number guessed and whether it is too high or too low.

Find the expected number $f(n)$ of calls needed to arrive at the number drawn, assuming that each guess is made at random from the values not already excluded.

Solution by José Heber Nieto, Universidad del Zulia, Maracaibo, Venezuela. The expected number of calls is $2(n+1)H_n/n - 3$, where $H_n = \sum_{k=1}^n 1/k$.

The probability that the first guess is correct is $1/n$. If the first guess is k and it is too high (an event with probability $(k-1)/n^2$), then we may expect $f(k-1)$ subsequent guesses. Similarly if the first guess is k and it is too low (an event with probability $(n-k)/n^2$), then we may expect $f(n-k)$ subsequent guesses. Thus

$$\begin{aligned} f(n) &= \frac{1}{n} + \sum_{k=2}^n \frac{k-1}{n^2} (1 + f(k-1)) + \sum_{k=1}^{n-1} \frac{n-k}{n^2} (1 + f(n-k)) \\ &= 1 + \frac{2}{n^2} \sum_{k=1}^{n-1} kf(k). \end{aligned}$$

From this we obtain

$$\begin{aligned} n^2 f(n) &= (n-1)^2 f(n-1) = \left(n^2 + 2 \sum_{k=1}^{n-1} kf(k) \right) - \left((n-1)^2 + 2 \sum_{k=1}^{n-2} kf(k) \right) \\ &= 2n - 1 + 2(n-1)f(n-1) \end{aligned}$$

and hence

$$f(n) = \frac{n^2 - 1}{n^2} f(n-1) + \frac{2n-1}{n^2}.$$

To solve this recurrence, let $g(n) = nf(n)/(n+1)$. Then $g(n)$ satisfies

$$g(n) = g(n-1) + \frac{2n-1}{n(n+1)}$$

with $g(1) = f(1)/2 = 1/2$. Thus

$$g(n) = \sum_{k=1}^n \frac{2k-1}{k(k+1)} = \sum_{k=1}^n \left(\frac{3}{k+1} - \frac{1}{k} \right) = 2H_n - \frac{3n}{n+1}$$

and therefore

$$f(n) = \frac{(n+1)g(n)}{n} = \frac{2(n+1)H_n}{n} - 3.$$

Editorial comment. Karl Hinderer pointed out that this problem was solved in E. Wong, "A Linear Search Problem", *SIAM Review*, 6 (1964), 168–174. Several solvers observed that the expected number of calls would be smaller if the callers adopted the strategy of guessing a number near the midpoint of the range of values not already excluded.

Several solvers also examined the asymptotic behavior of $f(n)$. The most common observation was that $f(n) \sim 2 \ln n$. Robert A. Agnew got a finer approximation, citing the result on harmonic numbers found in R. L. Graham, D. E. Knuth & O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1989, formula

6.66, p. 264.

$$H_n = \ln n + \gamma + \frac{1}{2n} + \frac{1}{12n^2} + O\left(\frac{1}{n^4}\right)$$

where $\gamma = .5772\dots$ is Euler's constant.

Solved by 26 readers (including those cited) and the proposer.

Counting Triangles in a Square

E 3450 [1991, 553]. *Proposed by Douglas Bowman (student), University of California at Los Angeles.*

Let $T(n)$ be the number of triangles lying in the subset $[0, n] \times [0, n]$ of the Cartesian plane whose sides lie on lines of slope 0, ∞ , 1, or -1 passing through points with integer coordinates. Derive a closed formula for $T(n)$.

Solution by Nasha Komanda, Central Michigan University, Mt. Pleasant, MI. The number of triangles is $\lfloor 3n^3 + \frac{9}{2}n^2 + n \rfloor$, which can also be written as $n(n+1)(3n+1) + \lfloor n^2/2 \rfloor$. For $k = 1, 2, \dots, n$, let $X(n, k)$ and $Y(n, k)$ be the sets of triangles in $[0, n] \times [0, n]$ with vertices of the form $\{(a, b), (a+k, b), (a, b+k)\}$ and $\{(a, b), (a+k, b), (a+k/2, b+k/2)\}$, respectively, where a and b are integers. Then $T(n) = 4\sum_{k=1}^n (|X(n, k)| + |Y(n, k)|)$. Since $0 \leq a \leq n-k$ and $0 \leq b \leq n-k$ for any triangle in $X(n, k)$, we have $|X(n, k)| = (n-k+1)^2$. Since $0 \leq a \leq n-k$ and $0 \leq b \leq \lfloor n-k/2 \rfloor$ for any triangle in $Y(n, k)$, we have $|Y(n, k)| = (n-k+1)(\lfloor n-k/2 \rfloor + 1)$.

Breaking the summation over $Y(n, k)$ into two parts according to the parity of k , we have

$$\begin{aligned} \frac{T(n)}{4} &= \sum_{k=1}^n (n-k+1)^2 + \sum_{j=1}^{\lfloor n/2 \rfloor} (n-2j+1)(n-j+1) \\ &\quad + \sum_{j=1}^{\lfloor n/2 \rfloor} (n-2j+2)(n-j+1). \end{aligned}$$

Now apply the formulas $\sum_{i=1}^m i = m(m+1)/2$ and $\sum_{i=1}^m i^2 = m(m+1)(2m+1)/6$ to obtain the answers that were claimed.

Editorial comment. Many readers erred in assuming additional conditions on the triangles. Richard K. Guy pointed out that this problem appeared in *Crux Mathematicorum* in 1981, having been translated from a Chinese New Year celebration in the *Scientific Daily* of Beijing, dated February 8, 1980.

Solved also by D. Beckwith, R. J. Chapman (U.K.), M. P. Eisner, J. Fukuta (Japan), J. H. Nieto (Venezuela), R. A. Winston, Anchorage Math Solutions Group, National Security Agency Problems Group, and the proposer. In addition, 14 incorrect solutions were received.

Diagonals are Hard to Measure

6662 [1991, 559]. *Proposed by F. S. Cater and John Erdman, Portland State University, Oregon.*

(a) Let I be the unit interval $[0, 1]$ and let $I \times I$ be the unit square. Let \mathcal{A} denote the smallest σ -algebra of subsets of $I \times I$ containing all rectangles of the

form $U \times V$ where either U or $I \setminus U$ is a first category set, and either V or $I \setminus V$ is a first category set. Prove that the diagonal $D = \{(x, x): x \in I\}$ does not lie in \mathcal{A} .

(b) Is this true when “first category set” is replaced by “set of measure zero”?

(c) Let a and b be cardinal numbers such that $a > b \geq \aleph_0$, and S be a set with cardinality $|S| = a$. Let \mathcal{B} denote the smallest σ -algebra of subsets of $S \times S$ containing all rectangles of the form $U \times V$ where either U or $S \setminus U$ has cardinality $\leq b$, and either V or $I \setminus V$ has cardinality $\leq b$. Prove that the diagonal $D = \{(x, x): x \in S\}$ does not lie in \mathcal{B} .

Solution I by Kenneth Schilling, University of Michigan, Flint, MI. Since there is no difficulty in proving the result when the countable unions required for a σ -algebra are generalized to unions of families indexed by a transfinite number κ , that generality will be used. We also use A_x for $\{y \in Y: (x, y) \in A\}$. All three parts of the problem are contained in the following proposition:

Proposition. *Let X and Y be sets, and let \mathcal{T} be a κ -ideal on X . Let \mathcal{A} be the smallest κ -algebra of subsets of $X \times Y$ containing all rectangles of the form $U \times V$, where $U \in \mathcal{T}$ or $X \setminus U \in \mathcal{T}$, and V is any subset of Y . Then for all $A \in \mathcal{A}$, there exists $T \subset Y$ such that*

$$\{x \in X: A_x \notin T\} \in \mathcal{T}. \quad (*)$$

Proof: The proof is by induction on the formation of the κ -algebra \mathcal{A} . Rectangles $U \times V$ with U or $X \setminus U \in \mathcal{T}$ clearly satisfy $(*)$. For the induction steps, note that if $\{x \in X: A_x \neq T\} \in \mathcal{T}$, then

$$\{x \in X: ((X \times Y) \setminus A)_x \neq (Y \setminus T)\} = \{x \in X: A_x \neq T\} \in \mathcal{T},$$

and if $\{x \in X: (A_\xi)_x \neq T_\xi\} = I_\xi \in \mathcal{T}$ for $\xi < \kappa$, then

$$\left\{x \in X: \left(\bigcap_{\xi < \kappa} A_\xi\right)_x \neq \bigcap_{\xi < \kappa} T_\xi\right\} \subset \bigcup_{\xi < \kappa} I_\xi \in \mathcal{T}$$

and the proof of the proposition is complete.

Since the diagonal D does not satisfy $(*)$ when \mathcal{T} is any of the ideals in (a), (b), or (c), the problem is solved. In addition, we can drop the condition on V in each part, and in part (c) enlarge \mathcal{B} to a b -ideal.

Solution II by J. Wengenroth (student), Universität Trier, Trier, Germany. In part (b) we assume only that the measure is such that single points have measure zero, but the whole interval does not. Then, all three parts of the problem follow from

Proposition. *Let S be any set, $\mathcal{B} \subseteq 2^S$ a σ -algebra (possibly $\mathcal{B} = 2^S$) and $\mathcal{C} \subset \mathcal{B}$ a system of sets with the following properties:*

- (1) $\{x\} \in \mathcal{C}$ for all $x \in S$,
- (2) if $V \in \mathcal{C}$ and $U \in \mathcal{B}$ with $U \subset V$ then $U \in \mathcal{C}$,
- (3) $A_n \in \mathcal{C}$ for $n \in \mathbb{N}$ implies $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{C}$,
- (4) $S \notin \mathcal{C}$.

For $\mathcal{M} = \{V \in \mathcal{B}: V \in \mathcal{C} \text{ or } S \setminus V \in \mathcal{C}\}$ and $\mathcal{A} = \sigma(\mathcal{M} \times \mathcal{M})$, the σ -algebra generated by $\mathcal{M} \times \mathcal{M}$, the diagonal $D = \{(x, x): x \in S\}$ does not lie in \mathcal{A} .

Proof: It is easily seen that \mathcal{M} is a σ -algebra over S . We define a measure ν on \mathcal{M} by

$$\nu(V) = \begin{cases} 1 & \text{if } V \in \mathcal{E} \\ 0 & \text{if } S \setminus V \in \mathcal{E}. \end{cases}$$

Properties (3) and (4) ensure that ν is σ -additive (because for a given sequence $\langle A_n \rangle$ of disjoint sets in \mathcal{M} there is at most one with $S \setminus A_n \in \mathcal{E}$). Let μ denote the product measure $\nu \otimes \nu$ on $\mathcal{A} = \mathcal{M} \otimes \mathcal{M}$ and μ^* the corresponding outer measure.

Suppose now $D \in \mathcal{A}$. Applying Fubini's theorem we see with (1)

$$\begin{aligned} \mu^*(D) &= \mu(D) = \int 1_D d\nu \otimes \nu \\ &= \int \nu(\{y \in S : (x, y) \in D\}) d\nu(x) \\ &= \int \nu(\{x\}) d\nu(x) = 0. \end{aligned}$$

On the other hand we can approximate $\mu^*(D)$ by sets of the form $\sum_{i=1}^n A_i$ (disjoint union) where A_i are elements of the semiring $\mathcal{M} \otimes \mathcal{M}$ (cf. M. M. Rao, *Measure Theory and Integration*, Wiley 1987, 2.2, Theorem 10), this means

$$\mu^*(D) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A_n \in \mathcal{M} \times \mathcal{M} \quad D \subset \bigcup_{n \in \mathbb{N}} A_n \right\}.$$

Since $\mu^*(D) = 0$ there are sets $A_n = B_n \times C_n \in \mathcal{M} \times \mathcal{M}$, $n \in \mathbb{N}$, with $D \subset \bigcup_{n \in \mathbb{N}} A_n$ and $\sum_{n=1}^{\infty} \mu(B_n \times C_n) < \frac{1}{2}$. Hence,

$$\mu(B_n \times C_n) = \nu(B_n) \cdot \nu(C_n) = 0 \quad \text{for all } n \in \mathbb{N}.$$

By the definition of D , for all $x \in S$ there is a $n \in \mathbb{N}$ such that $(x, x) \in B_n \times C_n$, that is $x \in B_n \cap C_n$. This implies

$$\begin{aligned} 1 = \nu(S) &= \nu\left(\bigcup_{n \in \mathbb{N}} B_n \cap C_n\right) \\ &\leq \sum_{n=1}^{\infty} \nu(B_n \cap C_n) \\ &\leq \sum_{n=1}^{\infty} \min(\nu(B_n), \nu(C_n)) = 0, \end{aligned}$$

a contradiction.

Solved also by E. Coplakova & K. P. Hart (The Netherlands), N. S. Feldman (student), I. Kastanas, A. Szymanski, and the proposers.

Another Stirling Identity

E 3451 [1991, 645]. Proposed by M. A. Khan, Lucknow, India.

Prove that for any positive integers m, n, N we have

$$\sum_{j=1}^{\min(n, N)} S(n, j) \frac{(N-j)^m}{(N-j)!} = \sum_{i=1}^{\min(m, N)} S(m, i) \frac{(N-i)^n}{(N-i)!}$$

where $S(n, j)$ is the Stirling number of the second kind given by

$$S(n, j) = \frac{1}{j!} \sum_{k=1}^j (-1)^{j-k} \binom{j}{k} k^n.$$

Solution by David Magagnosc, Drexel University, Philadelphia, PA. The Stirling number $S(n, j)$ counts the partition of an n element set into j nonempty parts. Letting $[n] = \{1, \dots, n\}$, the number of functions from $[n]$ onto $[j]$ is exactly $j!S(n, j)$.

Multiplying both sides of the proposed identity by $N!$ and rearranging, we find that we are asked to prove that

$$\sum_{j=1}^{\min(n, N)} \binom{N}{j} j! S(n, j) (N-j)^m = \sum_{i=1}^{\min(m, N)} \binom{N}{i} i! S(m, i) (N-i)^n.$$

These are equal because both sides count the pairs of functions (f, g) such that

$$f: [n] \rightarrow [N], \quad g: [m] \rightarrow [N],$$

and the images of f and g are disjoint. The first sum counts these pairs according to the size of the image of f : for a fixed size j , there are $\binom{N}{j}$ ways to select j elements to form the image, $j!S(n, j)$ ways to select f with that image, and then $(N-j)^m$ ways to select an arbitrary function g with range disjoint from that of f . Similarly, the second sum counts the pairs according to the size of the range of g .

Solved also by J. C. Binz (Switzerland), E. R. Canfield, R. J. Chapman (U.K.), W. Y. C. Chen, P. Čížek (student, Czechoslovakia), M. Dindos (Czechoslovakia), P. Haukkanen (Finland), A. A. Jagers (The Netherlands), I. Kastanas, K. S. Kedlaya (student), O. P. Lossers (The Netherlands), R. Martin (student), J. H. Nieto (Venezuela), A. Nijenhuis, J. Stirling, V. Strehl (Germany), E. Suárez (Spain), R. N. Will, Anchorage Math Solutions Group, and the proposer.

Collaborating editors: David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pfeifer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, Edward T. H. Wang, and William E. Watkins.

ANSWER to picture puzzle:

(on page 282)

Max Zorn who still lives and works at Indiana University.

The American Mathematical Monthly



Volume 100, Number 4 / APRIL 1993



AN OFFICIAL PUBLICATION OF THE MATHEMATICAL ASSOCIATION OF AMERICA

NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

Cover: Napier's portrait titled "Painting in Scotland, 1570 – 1650." The painting was originally in the possession of Margeret, Baroness of Napier who gave it to the University of Edinburgh, where it now hangs.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International,
Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 4 / APRIL 1993
(ISSN 0002-9890)



Contents

ARTICLES

- Taxicabs and Sums of Two Cubes / JOSEPH H. SILVERMAN 331
The Evil Twin Strategy for a Football Pool / JOSEPH DESTEFANO,
PETER DOYLE, and J. LAURIE SNELL 341
Six, Lies, and Calculators / R. M. CORLESS 344
What Is a Napierian Logarithm? / RAYMOND AYOUB 351
The Tyranny of Tests / PETER HILTON 365
A Really Trivial Proof of the Lucas-Lehmer Test / J. W. BRUCE 370
Pascal's Matrices / GREGORY S. CALL and
DANIEL J. VELLEMAN 372
Symmetries of the Cube and Outer Automorphisms of S_6 /
THOMAS A. FOURNELLE 377
Isogonal Configurations / TIMOTHY A. MURDOCH 381
-

FEATURES

- COMMENTS 330
NOTES 385
PICTURE PUZZLE 393
THE AUTHORS 394
LETTERS 396
UNSOLVED PROBLEMS 398
Are There only Finitely Many Binomial Coefficients with Positive
Deficiency?
PROBLEMS AND SOLUTIONS 400
REVIEWS
 Journey into Geometries by Marta Sved; *Roads to Geometry*
 by Edward C. Wallace and Stephen F. West /
 WILLIAM E. FENTON 411
 Polyominoes: A Guide to Puzzles and Problems in Tiling
 by George Martin / VICTOR G. FESER 412
TELEGRAPHIC REVIEWS 417

Taxicabs and Sums of Two Cubes

Joseph H. Silverman*

Our story begins in 1913, when the distinguished British mathematician G. H. Hardy received a bulky envelope from India full of page after page of equations. Every famous mathematician periodically receives letters from cranks who claim to have proven the most wonderous results. Sometimes the proofs, incorrect or incoherent, are included. At other times the writer solicits a reward in return for revealing his discoveries. Now this letter to Hardy, which was from a poor clerk in Madras by the name of Ramanujan, was filled with equations, all given without any sort of proof. Some of the formulas were well-known, mere exercises; while many of the others looked preposterous to Hardy's trained eye.

Who would have blamed Hardy if he had returned this missive to the sender, unread? And in fact, Ramanujan had previously sent his results to two other British mathematicians, each of whom had done just that! But instead Hardy gave some thought to these "wild theorems. Theorems such as he had never seen before, nor imagined."¹ And together, he and J. E. Littlewood, another eminent mathematician with whom Hardy often worked, succeeded in proving some of Ramanujan's amazing identities. At this point Hardy realized that this letter was from a true mathematical genius, and he became determined that Ramanujan should come to England to pursue his mathematical researches. Using travel money provided by Hardy's college, Ramanujan arrived in 1914. Over the next several years he continued to produce and publish highly original material, and he also collaborated with Hardy on a number of outstanding papers.

In 1918, at the age of 30, Ramanujan was elected a Fellow of the Royal Society and also of Trinity College, both signal honors which he richly deserved. Unfortunately, in the colder climate of England he contracted tuberculosis. He returned to his native Madras and died, in 1920, at the age of 33.

During all of Ramanujan's life, he considered numbers to be his personal friends. To illustrate, Hardy tells the story of how one day he visited Ramanujan in the hospital. At a loss for something to say, Hardy remarked that he had arrived at the hospital in taxicab number 1729. "It seemed to me," he continued, "a rather dull number." To which Ramanujan replied "No, Hardy! It is a very interesting number. It is the smallest number expressible as a sum of two cubes in two different ways."²

$$1729 = 1^3 + 12^3 = 9^3 + 10^3.$$

*This article is an expanded version of talks given at M.I.T. and Brown University

¹[10], page 32

²[10], page 37

Hardy then asked for the smallest number which is a sum of two *fourth* powers in two different ways, but Ramanujan did not happen to know.³

Rather than studying sums of higher powers, we will instead concern ourselves with another question that Hardy could just as easily have asked, namely for the smallest number which is a sum of two cubes in three (or more) distinct ways. In this case the answer is given in [3],

$$87,539,319 = 436^3 + 167^3 = 423^3 + 228^3 = 414^3 + 255^3,$$

although if one is willing to allow both positive and negative integers there is the much smaller solution

$$4104 = 16^3 + 2^3 = 15^3 + 9^3 = (-12)^3 + 18^3.$$

In this note we will be taking a leisurely number-theoretic stroll centered around the problem of writing numbers as sums of two cubes, specifically the search for integers with many such representations. It will be some time before we actually return to this specific question, but along the way we will view some beautiful mathematics which illustrates some of the interplay that can occur between geometry, algebra, and number theory.

We will thus be looking at solutions of the equation

$$X^3 + Y^3 = A.$$

What can be said about such solutions? First we have the elementary result that if A is a non-zero integer, then there are only finitely many solutions in integers X and Y . Of course, if X and Y are restricted to be positive integers, then this is obvious. But in any case we can use the factorization

$$A = X^3 + Y^3 = (X + Y)(X^2 - XY + Y^2)$$

to see that A must factor as $A = BC$ in such a way that

$$B = X + Y \quad \text{and} \quad C = X^2 - XY + Y^2.$$

Now there are only finitely many ways of factoring A as $A = BC$, and for each such factorization we substitute the first equation (i.e. $Y = B - X$) into the second to obtain

$$X^2 - X(B - X) + (B - X)^2 = C.$$

Thus each factorization of A yields at most two values for X , each of which gives one value for $Y = B - X$. Therefore there are only finitely many integer solutions (X, Y) .⁴

[Aside: This proof that the equation $X^3 + Y^3 = A$ has only finitely many solutions in integers depends heavily on the factorization of the polynomial $X^3 + Y^3$. It is similarly true, but quite difficult to prove, that an equation like $X^3 + 2Y^3 = A$ has only finitely many integer solutions. This was first proven by A. Thue in 1909 [11]. By way of contrast, note that the equation $X^2 - 2Y^2 = 1$ has infinitely many solutions.]

³The answer, $635,318,657 = 59^4 + 158^4 = 133^4 + 134^4$, appears to have been discovered by Euler. And it does not seem to be known if there are any numbers which are a sum of two fifth powers in two different ways.

⁴Exercise: Show that any solution in integers satisfies $\max\{|X|, |Y|\} \leq 2\sqrt{|A|/3}$.

As is generally the case in mathematics, when a person wants to work on a difficult mathematical problem, it is nearly always best to start with a related easier problem. Then, step by step, one approaches the original goal. In our case we are confronted with the equation

$$X^3 + Y^3 = A,$$

and a natural first question is to ask for the solutions in real numbers. In other words, what does the graph of this curve look like? Since $dY/dX = -(X/Y)^2$, the graph is always falling. Further, it is symmetric about the line $Y = X$ and has the line $Y = -X$ as an asymptote. With this information it is easy to sketch the graph, which is illustrated in Figure 1. We will denote the resulting curve by C .

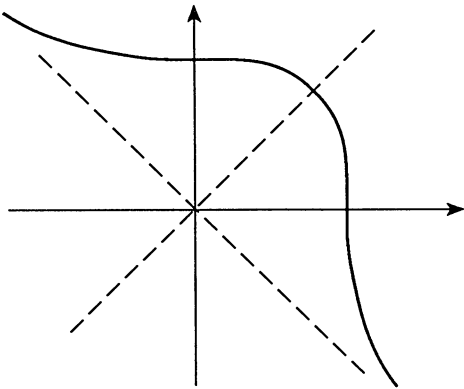


Figure 1. The curve $C: X^3 + Y^3 = A$

If (X, Y) is a point on this curve C , then so is the reflected point (Y, X) . But there is another, less obvious, way to produce new points on the curve C which will be very important for the sequel. Thus suppose that $P = (X_1, Y_1)$ and $Q = (X_2, Y_2)$ are two (distinct) points on C , and let L be the line connecting P to Q . Then L will (usually) intersect C at exactly one other point. To see this, suppose that L is given by the equation $Y = mX + b$. Then substituting the equation of L into the equation of C gives the cubic equation

$$X^3 + (mX + b)^3 = A.$$

By assumption, this equation already has the solutions X_1 and X_2 , since P and Q lie on the intersection $C \cap L$, so the cubic equation has exactly one other solution X_3 . (We run into a problem if $m = -1$, but we'll deal with that later.) Then letting $Y_3 = mX_3 + b$, we have produced a new point $R = (X_3, Y_3)$ on C . Further, even if $P = Q$, the same procedure will work provided we take L to be the tangent line to C at P . Thus given any two points P and Q on C , we have produced a third point R . Finally we define an “addition law” on C by setting

$$P + Q = (\text{reflection of } R \text{ about the line } Y = X).$$

In other words, if $R = (X_3, Y_3)$, then $P + Q = (Y_3, X_3)$. (See Figure 2.) The reason we define $P + Q$ with this reflection will soon become clear, but first we have to deal with the pesky problem that sometimes our procedure fails to yield a third point.

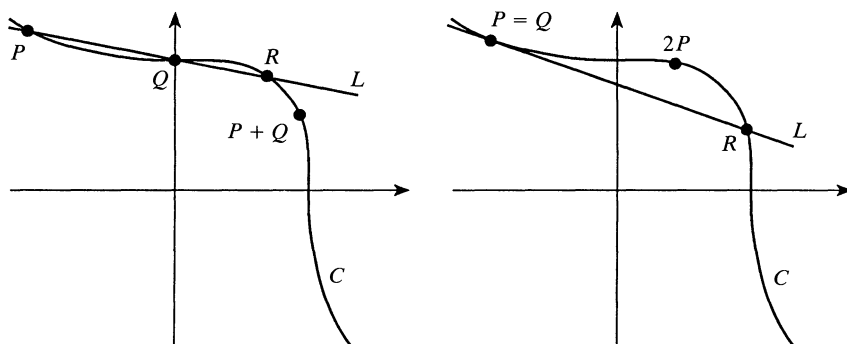


Figure 2. The addition law on the curve $C: X^3 + Y^3 = A$

This problem arises whenever the line connecting P and Q has slope -1 , and thus is parallel to the asymptote $Y = -X$. We solve the problem in cavalier fashion; since there is no actual third intersection point we create one by fiat. To be precise, we take the XY -plane and add an extra point, which we will denote by \mathcal{O} . This point has the property that the lines going through \mathcal{O} are exactly the lines with slope -1 . Further, if P is a point in the XY -plane, then the unique line through P and \mathcal{O} is the line through P having slope -1 . Having done this, we are now in the happy position of being able to assert that given any two points P and Q on C , the above procedure will yield a unique third point R on C , so we are able to define $P + Q$ for any P and Q . (By definition, we set $\mathcal{O} + \mathcal{O} = \mathcal{O}$.)

[*Aside.* Some readers will doubtless recognize that what we have done is start the construction of the real projective plane. The projective plane consists of the usual XY -plane together with one point for each direction. In our case, we only needed the direction with slope -1 . The projective plane has the agreeable property that any two lines, even “parallel” ones, intersect at exactly one point. This is true because if two lines are parallel, then they intersect at the point corresponding to their common direction.]

The justification for our use of the symbol “+” is now this:

The addition law $P + Q$ described above makes the points of C into an abelian group.

To be more precise, if $P = (X, Y)$ is a point on C , we define its inverse to be its reflection, $-P = (Y, X)$. Then for all points P, Q, R on C we have

$$P + \mathcal{O} = \mathcal{O} + P = P \quad (\text{identity})$$

$$P + (-P) = \mathcal{O} \quad (\text{inverse})$$

$$P + Q = Q + P \quad (\text{commutativity})$$

$$(P + Q) + R = P + (Q + R) \quad (\text{associativity}).$$

All of these properties are quite easy to check except for the associativity, which we will return to in a moment.

Note that the addition procedure outlined above is entirely mechanical. We can write down explicit formulas for the sum of two points, although there are a number of special cases. For example, if $P = (X, Y)$, then

$$P + P = 2P = \left(\frac{Y(X^3 + A)}{Y^3 - X^3}, \frac{X(Y^3 + A)}{X^3 - Y^3} \right).$$

The formula for the sum of two distinct points $P_1 = (X_1, Y_1)$ and $P_2 = (X_2, Y_2)$ is somewhat more complicated:

$$\begin{aligned} P_1 + P_2 \\ = \left(\frac{A(Y_1 - Y_2) - X_1X_2(Y_1X_2 - Y_2X_1)}{X_1X_2(X_1 - X_2) + Y_1Y_2(Y_1 - Y_2)}, \frac{A(X_1 - X_2) - Y_1Y_2(X_1Y_2 - X_2Y_1)}{X_1X_2(X_1 - X_2) + Y_1Y_2(Y_1 - Y_2)} \right). \end{aligned}$$

And now that we are in possession of these formulas, verification of the associative law is a tedious, but straightforward, task.

Let us now look at an example, and what better choice than Ramanujan's equation

$$X^3 + Y^3 = 1729.$$

We already know two interesting points on this curve, namely

$$P = (1, 12) \quad \text{and} \quad Q = (9, 10).$$

Using the addition law, we can easily compute some more points, such as

$$\begin{aligned} P + Q &= \left(\frac{46}{3}, \frac{-37}{3} \right), & P - Q &= \left(\frac{453}{56}, \frac{-397}{56} \right), & 2P &= \left(\frac{20760}{1727}, \frac{-3457}{1727} \right), \\ 2Q &= \left(\frac{24580}{271}, \frac{-24561}{271} \right), & 3P &= \left(\frac{-5150812031}{107557668}, \frac{5177701439}{107557668} \right). \end{aligned}$$

As the discerning reader will notice, the numbers produced seem to grow with frightening rapidity. But of far more importance is the fact that although we have not produced any new integer solutions, all of the new solutions are at least in rational numbers.

This is a consequence of the fact that the addition law on C is given by rational functions. That is, the coordinates of $P + Q$ are given by quotients of polynomials in the coordinates of P and Q . Thus if the coordinates of P and Q are rational numbers, then so are the coordinates of $P + Q$. (If this is not clear, look for example at the formula for $2P$ given above. If A , X , and Y are all rational numbers, then the formula shows that the coordinates of $2P$ are also rational.) Let us define

$$C(\mathbb{Q}) = \{(X, Y) : X \text{ and } Y \text{ are rational numbers and } X^3 + Y^3 = A\} \cup \{\mathcal{O}\}.$$

Here \mathbb{Q} is the usual symbol for the field of rational numbers. Then the remarks given above provide a proof of the following fact, first noted by Poincaré around 1900 [6]:

The set $C(\mathbb{Q})$ is a subgroup of C . In other words, $C(\mathbb{Q})$ supplied with the addition law from C becomes a group in its own right.

The sum of two cubes “Hall of Fame”

The following material gives the smallest integer I know with the indicated property. It has been collected from various sources, including [3] and [12]. Note that we are counting $X^3 + Y^3$ and $Y^3 + X^3$ as being the same representation.

Positive integers, 2 representations, cube-free

$$1729 = 1^3 + 12^3 = 9^3 + 10^3 = 7 \cdot 13 \cdot 19$$

Any integers, 3 representations, not cube-free

$$4104 = 16^3 + 2^3 = 15^3 + 9^3 = (-12)^3 + 18^3 = 2^3 \cdot 3^3 \cdot 19$$

Any integers, 3 representations, cube-free

$$3,242,197 = 141^3 + 76^3 = 138^3 + 85^3 = (-171)^3 + 202^3 = 7 \cdot 31 \cdot 67 \cdot 223$$

Positive integers, 3 representations, not cube-free

$$\begin{aligned} 87,539,319 &= 436^3 + 167^3 = 423^3 + 228^3 = 414^3 + 255^3 \\ &= 3^3 \cdot 7 \cdot 31 \cdot 67 \cdot 223 \end{aligned}$$

Positive integers, 3 representations, cube-free

$$\begin{aligned} 15,170,835,645 &= 517^3 + 2468^3 = 709^3 + 2456^3 = 1733^3 + 2152^3 \\ &= 3^2 \cdot 5 \cdot 7 \cdot 31 \cdot 37 \cdot 199 \cdot 211 \end{aligned}$$

Any integers, 4 representations, not cube-free

$$\begin{aligned} 42,549,416 &= 348^3 + 74^3 = 282^3 + 272^3 \\ &= (-2662)^3 + 2664^3 = (-475)^3 + 531^3 \\ &= 2^3 \cdot 7 \cdot 13 \cdot 211 \cdot 277 \end{aligned}$$

Positive integers, 4 representations, not cube-free

$$\begin{aligned} 26,059,452,841,000 &= 29620^3 + 4170^3 = 28810^3 + 12900^3 \\ &= 28423^3 + 14577^3 = 24940^3 + 21930^3 \\ &= 2^3 \cdot 5^3 \cdot 31 \cdot 43^3 \cdot 97 \cdot 109 \end{aligned}$$

Any integers, 4 representations, cube-free

Unknown!

Any integers, 5 representations, not cube-free

$$\begin{aligned} 1,148,834,232 &= 1044^3 + 222^3 = 920^3 + 718^3 \\ &= 846^3 + 816^3 = (-7986)^3 + 7992^3 = (-1425)^3 + 1593^3 \\ &= 2^3 \cdot 3^3 \cdot 7 \cdot 13 \cdot 211 \cdot 277 \end{aligned}$$

We have come a long way in our study of the solutions of the equation $X^3 + Y^3 = A$. What we have found is that the set of solutions in rational numbers becomes, in a very natural way, an abelian group. So if we can say something significant about this group, then we might feel that at last we have some understanding about this set of rational solutions. An answer to this problem is provided by one of the most celebrated theorems of the twentieth century. Aside from its intrinsic interest, this result has been the starting point for much of the study of Diophantine equations over the past 70 years. The theorem, as we state it, was first proven by L. J. Mordell in 1922 [5]. It was subsequently vastly generalized

by A. Weil in his 1928 thesis, and so is usually called the *Mordell-Weil Theorem*:

There exists a finite set of points P_1, \dots, P_r in $C(\mathbb{Q})$ so that every point in $C(\mathbb{Q})$ can be obtained from P_1, \dots, P_r by addition and subtraction. In fancier language, the group $C(\mathbb{Q})$ is finitely generated. The points P_1, \dots, P_r are called generators for $C(\mathbb{Q})$.

This seems fairly satisfactory. Even though our equation may have infinitely many rational solutions, they can all be obtained by starting with some finite subset and applying a completely mechanical procedure. For example, on the curve

$$X^3 + Y^3 = 7,$$

every rational point is a multiple of the single generating point $(2, -1)$. Similarly, it is probably true that every rational point on Ramanujan's curve has the form $nP + mQ$, where $P = (1, 12)$, $Q = (9, 10)$, and n and m are allowed to range over all integers, although I am not aware that anyone has verified this probable fact.

It may come as a surprise, then, to learn that the Mordell-Weil Theorem is far less satisfactory than it appears. This is due to the fact that it is not *effective*. What this means is that currently we do not have an algorithm which will determine, for every value of A , a set of generators P_1, \dots, P_r for $C(\mathbb{Q})$. In fact, there is not even a procedure for determining exactly how many generators are needed, although it is possible to give a rather coarse upper bound. This problem of making the Mordell-Weil Theorem effective is one of the major outstanding problems in the subject.

Another open problem concerns the number of generators needed. As mentioned above, the curve $X^3 + Y^3 = 7$ requires only one generator, while Ramanujan's curve $X^3 + Y^3 = 1729$ needs at least two. The question is whether there are curves which require a large number of generators. More precisely, is it true that for every integer r there is some value of A so that the rational points on the curve $X^3 + Y^3 = A$ require at least r generators?

We now return to our original problem, namely the study of *integer* solutions to the equation $X^3 + Y^3 = A$. Specifically, we seek values of A for which this equation has many solutions. We have observed that the equation

$$X^3 + Y^3 = 7$$

has the solution $P = (2, -1)$, and by taking multiples of P we can produce a sequence of points

$$P = (2, -1), \quad 2P = \left(\frac{5}{3}, \frac{4}{3}\right), \quad 3P = \left(\frac{-17}{38}, \frac{73}{38}\right), \quad 4P = \left(\frac{-1256}{183}, \frac{1265}{183}\right), \dots$$

Further, it is true (but moderately difficult to prove) that this sequence of points $P, 2P, 3P, \dots$ never repeats. Each point in the sequence has rational coordinates, so we can write nP in the form

$$nP = \overbrace{P + P + \dots + P}^{n\text{-terms}} = \left(\frac{a_n}{d_n}, \frac{b_n}{d_n}\right),$$

where a_n , b_n , and d_n are integers. We are going to take the first N of these rational solutions and multiply our original equation by a large integer so as to clear the denominators of all of them. Thus let

$$B = d_1 d_2 \cdots d_N.$$

Then the equation

$$X^3 + Y^3 = 7B^3$$

has at least N solutions in integers, namely

$$\left(\frac{a_n B}{d_n}, \frac{b_n B}{d_n} \right), \quad 1 \leq n \leq N.$$

(Actually $2N$ solutions, since we can always switch X and Y , but for simplicity we will generally count pairs of solutions (X, Y) and (Y, X) .) We have thus found the following answer to our original problem:

Given any integer N , there exists a positive integer A for which the equation $X^3 + Y^3 = A$ has at least N solutions in integers.

Of course, Ramanujan and Hardy were probably talking about solutions in positive integers; but with a bit more work one can show that infinitely many of the points $P, 2P, 3P, \dots$ have positive coordinates, so we even get positive solutions.

Naturally, it is of some interest to make this result quantitative, that is, to describe how large A must be for a given value of N . The following estimate is essentially due to K. Mahler [4], with the improved exponent appearing in [8]:

There is a constant $c > 0$ such that for infinitely many positive integers A , the number of positive integer solutions to the equation $X^3 + Y^3 = A$ exceeds $c\sqrt[3]{\log A}$.

In some sense we have now answered our original question. There are indeed integers which are expressible as a sum of two cubes in many different ways. But a nagging disquiet remains. We have not really produced a large number of intrinsically integral solutions. Rather, we have cleared the denominators from a lot of rational solutions. In the solutions produced above, the X and Y coordinates will generally have a large common factor whose cube will divide A . What happens if we rule out this situation? The simplest way to do so is to restrict attention to integers A that are *cube-free*; that is, A should not be divisible by the cube of any integer greater than 1. This is a reasonable restriction since if D^3 divides A , then an integer solution (X, Y) to $X^3 + Y^3 = A$ really arises from the rational solution $(X/D, Y/D)$ to the “smaller” equation $X^3 + Y^3 = A/D^3$.

Notice Ramanujan’s example

$$1729 = 1^3 + 12^3 = 9^3 + 10^3 = 7 \cdot 13 \cdot 19$$

is cube-free. But the example with three representations given earlier,

$$87,539,319 = 436^3 + 167^3 = 423^3 + 228^3 = 414^3 + 255^3 = 3^3 \cdot 7 \cdot 31 \cdot 67 \cdot 223,$$

is not cube-free. As far as I have been able to determine, the smallest cube-free integer which can be expressed as a sum of two positive cubes in three distinct ways was unearthed by P. Vojta in 1983 [12]:

$$\begin{aligned} 15,170,835,645 &= 517^3 + 2468^3 = 709^3 + 2456^3 = 1733^3 + 2152^3 \\ &= 3^2 \cdot 5 \cdot 7 \cdot 31 \cdot 37 \cdot 199 \cdot 211. \end{aligned}$$

And now our problem has become so difficult that Vojta’s number holds the current record! There is no cube-free number known today which can be written as a sum of two positive cubes in four or more distinct ways.

We are going to conclude by describing a relationship between the two problems that we have been studying. The first problem was that of expressing a number as a sum of two *rational cubes*, and in that case we saw that all solutions arise from a finite generating set and speculated as to how large this generating set might be. The second problem was that of writing a cube-free number as a sum of two *integral cubes*, and here we saw that there are only finitely many solutions and we wondered how many solutions there could be. It is not a priori clear that these two problems are related, beyond the obvious fact that an integral solution is also a rational solution. In 1974 V. A. Dem'janenko stated that if there are a large number of integer solutions, then any generating set for the group of rational points must also be large, but his proof was incomplete. (See [2, page 140] for Lang's commentary on and generalization of Dem'janenko's conjecture.) The following more precise version of the conjecture was proven in 1982 [8]:

For each integer A , let $N(A)$ be the number of solutions in integers to the equation $X^3 + Y^3 = A$, and let $r(A)$ be the minimal number of rational points on this curve needed to generate the complete group of rational points (as in the Mordell-Weil Theorem). There is a constant $c > 1$ such that for every cube-free integer A ,

$$N(A) \leq c^{r(A)}.$$

Note that the requirement that A be cube-free is essential. For as we saw above, we can make $N(7B^3)$ as large as we want by choosing an appropriate value of B . On the other hand, $r(7B^3) = r(7)$ for every value of B , so an inequality of the form $N(A) \leq c^{r(A)}$ cannot be true if we allow arbitrary values of A . (To see that $r(7B^3) = r(7)$, note that the groups of rational points on the curves $X^3 + Y^3 = 7$ and $X^3 + Y^3 = 7B^3$ are the same via the map $(X, Y) \rightarrow (BX, BY)$.)

Final Remark. The cubic curve $X^3 + Y^3 = A$ is an example of what is called an *elliptic curve*. Those readers interested in learning more about the geometry, algebra, and number theory of elliptic curves might begin with [9] and continue with the references listed there. Elliptic curves and the related theory of elliptic functions appear frequently in areas as diverse as number theory, physics, computer science, and cryptography.

ACKNOWLEDGMENTS. I would like to thank Jeff Achter and Greg Call for their helpful suggestions.

REFERENCES

1. Dem'janenko, V. A.: On Tate height and the representation of a number by binary forms, *Math. USSR Isv.* **8**, 463–476 (1974).
2. Lang, S.: *Elliptic Curves: Diophantine Analysis*. Berlin: Springer-Verlag, 1978
3. Leech, J.: Some solutions of Diophantine equations. *Proc. Camb. Philos. Soc.* **53**, 778–780 (1957).
4. Mahler, K.: On the lattice points on curves of genus 1. *Proc. London Math. Soc.* **39**, 431–466 (1935).
5. Mordell, L. J.: On the rational solutions of the indeterminate equation of the third and fourth degree. *Proc. Camb. Math. Soc.* **21**, 431–466 (1922).
6. Poincaré, H.: Sur les propriétés arithmétiques des courbes algébriques. *J. de Liouville* **7**, 161–233 (1901).
7. Silverman, J. H.: Integer points and the rank of Thue elliptic curves. *Invent. Math.* **66**, 395–404 (1982).
8. Silverman, J. H.: Integer points on curves of genus 1. *J. London Math. Soc.* **28**, 1–7 (1983).

9. Silverman, J. H., Tate, J.: *Rational Points on Elliptic Curves*. New York: Springer-Verlag, 1992.
10. Snow, C. P., Foreword to: *A Mathematician's Apology*. by G. H. Hardy (1940), Cambridge: Cambridge University Press, 1967.
11. Thue, A.: Über Annäherungswerte algebraischer Zahlen. *J. reine ang. Math.* **135**, 284–305 (1909).
12. Vojta, P.: private communication, 1983.

Added in proof: Richard Guy has pointed out to the author that Rosenstiel, Dardis and Rosenstiel have recently found the non-cube-free example

$$\begin{aligned} 6963472309248 &= 2421^3 + 19083^3 = 5436^3 + 18948^3 + 10200^3 + 18072^3 \\ &= 13322^3 + 16630^3 \end{aligned}$$

which is smaller than the example in the above table, and which is in fact the smallest such example. See *Bull. Inst. Math. Appl.* 27(1991), 155–157.

Mathematics Department
Brown University
Providence, RI 02912 USA
jhs@gauss.math.brown.edu

Five Borromean Square Frames

William W. Chernoff

In [R. Brown and J. Robinson, Borromean circles, *Amer. Math. Monthly*, 99 (1992), 376–377], R. Brown and J. Robinson gave a Borromean arrangement of three square frames in \mathbb{R}^3 . In figure 1, we give a Borromean arrangement consisting of five square frames. In addition, we have constructed models consisting of seven hexagonal frames and of nine octagonal frames. A paper discussing Borromean arrangements in \mathbb{R}^3 of $2n + 1$ frames each with $2n$ sides is to appear.

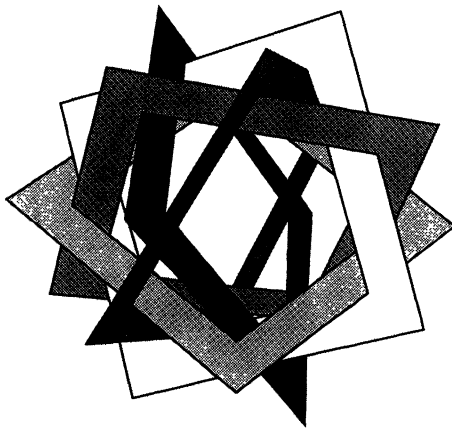


Fig. 1

Department of Mathematics and Statistics
 University of New Brunswick
 Fredericton, N.B.
 Canada E3B 5A3

The Evil Twin Strategy for a Football Pool

Joseph DeStefano, Peter Doyle and J. Laurie Snell

At Emmy's we have a weekly football pool in which each person puts \$1 in the pot and picks the winning team in about ten football games. The one who picks the most winners gets the entire pot. A point spread is provided, and the favored team must win by the number of points specified. The effect of the point spread is to make the probability of picking the winner of a particular game approximately $1/2$. Joe observed that most of us were just guessing and that some were even tossing a coin to decide which team to pick. He asked if he could contribute \$2 to the pot and submit two entries. Joe proposed to submit his choices and those of his evil twin. For each game, his evil twin picks the winner to be the team that Joe picks to lose. This request led to interesting moral, legal, and mathematical discussions and to the following problem.

Assume that all the participants in the football pool except Joe submit one entry and toss a fair coin to make their choices. Joe submits two entries: one for himself and one for his evil twin. For his choices, Joe also tosses a coin. For each game, Joe's evil twin picks as winner the team that Joe picks to lose. The player with the maximum score wins the entire pot. If there is more than one player with the maximum score, the winning players split the pot evenly. Is this a favorable game to Joe? That is, is his expected winning greater than \$2?

If Joe and his evil twin are competing only with Dana and there is only one game, the computation is easy. Either Joe or his evil twin will be correct. If Dana is wrong he wins nothing, and if he is correct he shares the \$3 pot with either Joe or his evil twin. Thus Dana's expected winning is

$$\frac{3}{2} \cdot \frac{1}{2} = \frac{3}{4}$$

and Joe's expected winning is

$$3 - \frac{3}{4} = \frac{9}{4} = 2.25.$$

In this trivial situation, Joe has an expected profit of 25 cents. To see if this strategy continues to pay off in a more realistic situation, Joe computed his expected profit while varying the number of players and the number of games. The result is shown in Figure 1.

From this computation, Joe noticed that, for a situation typical for our pool, with 10 games and 10 players, his advantage is still significant. In addition, he noticed the surprising fact that, for an odd number of games, the expected profit is independent of the number of games. We now prove this.

Assume that, in addition to Joe and his evil twin, there are m players tossing coins to choose the winners for $2n + 1$ games. Either Joe or his evil twin must

		Number of players							
		3	4	5	6	7	8	9	10
Number of Games	1	.250	.333	.344	.325	.297	.268	.241	.218
	2	.188	.250	.264	.258	.244	.229	.212	.197
	3	.250	.333	.344	.325	.297	.268	.241	.218
	4	.215	.286	.299	.288	.269	.248	.227	.208
	5	.250	.333	.344	.325	.297	.268	.241	.218
	6	.226	.301	.313	.300	.278	.254	.232	.212
	7	.250	.333	.344	.325	.297	.268	.241	.218
	8	.231	.308	.320	.306	.283	.258	.234	.213
	9	.250	.333	.344	.325	.297	.268	.241	.218
	10	.235	.313	.325	.310	.285	.260	.236	.214

Figure 1. Expected profit for Joe and his evil twin

have more than half correct. Assume that Joe does. Then Joe will be competing on an equal basis with the subset of other players who have more than half correct. Thus, if j of the other players get more than $1/2$ correct the expected payment to Joe is $(m + 2)/(j + 1)$. The number j is itself a chance quantity and can be viewed as the number of heads that turn up in m tosses of a coin. Thus, the expected return R for Joe and his evil twin is

$$R = \sum_{j=0}^m \frac{m+2}{j+1} \binom{m}{j} \frac{1}{2^m} = \frac{m+2}{m+1} \sum_{j=0}^m \binom{m+1}{j+1} \frac{1}{2^m} = \frac{2(m+2)}{m+1} \left(1 - \frac{1}{2^{m+1}}\right).$$

Note that the number of games never even enters into this argument.

Replacing m by x in our expression for R we obtain a corresponding continuous function. By plotting this function, or by elementary calculus, one can verify that for $x \geq 1$ its values are greater than 2. It increases to a maximum value 2.345 at $x = 2.693$ and then decreases approaching 2 as x goes to infinity. Note from our table that the most advantageous situation with an expected profit of .344 occurs with five players ($m = 3$).

A similar argument provides the expected return for the case of an even number of games, but the result is more complicated and not independent of the number of games. For a given number of players the expected return increases to the value for an odd number of games as the number of games tends to infinity.

The evil twin strategy poses many other interesting questions. Impressed by the above calculation, Prosser and Snell employ Joe's strategy with Prosser making the picks and Snell playing the role of the evil twin. They have done quite well, but Prosser seems to win more often than Snell. This suggests that Prosser might have a probability greater than $1/2$ of picking the winner of a game and raises the question how large this probability must be so that Prosser should reject Snell and play alone. The answer, for the case of 9 games, is that he has only to have a probability of .52 for guessing correctly the team that will win to be better off without the cooperation of Snell when the other players are guessing.

If Prosser and Snell are in collusion and Joe and his evil twin are also entering the pool how much better could these four players do by forming a coalition? More generally, what is the optimal strategy for a group of players that want to form a coalition? This question can be described geometrically as follows: A group of people are picking points at the corners of an n -dimensional cube, where n is the number of games, with the objective of getting as close as possible to the choice of a man from Mars, who picks a corner point at random. If m people make their choices at random how should r additional people choose their points to

maximize the probability that one of their points is the closest to the point picked by the man from Mars? This geometric interpretation also shows why Joe has an advantage. If you look at the case of three games, so that the points are picked on an ordinary cube, you will notice that Joe and his evil twin pick points on diagonally opposite corners and by so doing have more points near one of their two points than would be true typically for a pair of players who picked their points at random.

We thank our colleagues at Emmy's for many helpful discussions about how to participate in a football pool.

Joseph DeStefano and J. Laurie Snell:
Department of Mathematics and
Computer Science
Dartmouth College
Hanover, NH 03755

Peter Doyle:
Department of Mathematics
University of California, San Diego
LaJolla, CA 92093-0112

Trivia Mathematica

According to reports from the recent Columbus meetings of the Association and the Society, a new publication, *Trivia Mathematica*, is about to emerge from its chrysalis. At one of those profound meetings which last far into the night, learned savants sired and damned this new idea. Reports, though meager, place Brown and Princeton as focal points, with Cornell and M.I.T. as vortices of the hyperboles. Apparently without reMORSE, Princeton has added HURWITZ to those of divine Providence to FLOOD Cambridge and Ithaca with puns. No one has been judged best at this game, we understand, but doubtless by general consent they place WIENER-worst.

E.J.M.

—*American Mathematical Monthly*
47, (1940) p. 43.

Six, Lies, and Calculators

R. M. Corless

This article attempts to alleviate some dismal impressions that might be inadvertently left after reading Yves Nievergelt's otherwise excellent paper [1]. In particular, the present article shows that in a certain sense, the answers provided by the calculator are precisely as useful as the unique *exact* solution, for a very large class of problems, even when the calculator solutions are not at all close to the exact solution. It is further shown that even the exact answer by itself is usually *insufficient* for a complete understanding of the problem under consideration, and that some sort of sensitivity analysis (such as computing the condition number of the problem) is also required. It is also shown that unexpected behaviour of a computed solution can be extremely useful pedagogically.

The behaviour of the calculator is by no means unique. All other computers and software packages exhibit similarly useful quirks. Further, the behaviour is not restricted to the solution of linear equations, holding true also for function evaluation, rootfinding, the solution of differential equations, and many other practical problems. There is a strong connection with the theory of chaotic dynamics here, which may also be of interest.

SUMMARY OF Y. NIEVERGELT'S PAPER. Nievergelt's paper [1] reports what happens when you try to solve the linear system $A\mathbf{x} = \mathbf{b}$, given below, on the HP28S.

$$\begin{bmatrix} 888445 & 887112 \\ 887112 & 885781 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Several methods are tried, and only one produces the exact answer.

Professor Nievergelt correctly points out that the matrix A is very close to a singular matrix, and thus is extremely ill-conditioned, having a condition number greater than 10^{12} . A succinct and accurate algebraic definition of condition number is given in [1] (and hence is not repeated here), and Professor Nievergelt also points out that the definition, meaning, and use of the condition number are not sufficiently well known outside of the numerical analysis community, and should be taught more widely. I agree whole-heartedly with this. In fact this should be easy, because there is a strong connection between the condition number in the more general computational situation and differentials of first year calculus, and so in some sense the teaching of the "conditioning" of a problem is a simple application of part of mainstream calculus.

Of course there are many textbooks which discuss condition number in the context of numerical linear algebra and finding roots of polynomials (e.g. [2]), and others that discuss the condition number of function evaluation (e.g. [3]), and still others that do so for differential equations (e.g. [4]). This paper attempts an introductory overview.

ERRORS IN MODELLING, ERRORS IN DATA, AND ERRORS IN COMPUTATION. The key rationale for having to deal with condition numbers is *not* numerical analysis and the problem of computational error. A very large number of mathematical problems are derived from “real-world” origins, and contain both modelling error (e.g. neglected terms in the equations) and data or measurement error. These are unavoidable, while computational error is avoidable, at least in principle, if you wish to pay the price for doing exact arithmetic, for example using a symbolic manipulation package such as Maple [5].

It is a very useful feature of numerical analysis that the techniques used for monitoring the effects of computational errors can often be used to monitor the effects of modelling error and measurement error.

The basic principle is Wilkinson’s idea of *backward error analysis*: a good numerical method will give you the *exact* solution of a nearby problem. This very powerful idea reduces the study of computational errors to the study of modelling or measurement errors, *which we have to study anyway*.

This principle was first elucidated in the context of the solution of linear systems of equations and in the context of polynomial rootfinding. Instead of repeating details of these, I urge the reader to examine the technical and historical discussions in [2]. Examination of Figure 1 at this point may make the basic idea more clear.

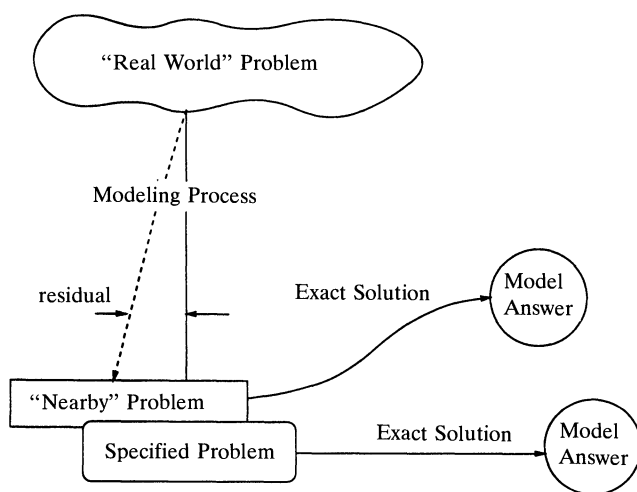


Figure 1. Modeling based on a nearby problem.

The fundamental point is that we get insight from knowing exact solutions—that is, from knowing both the question and the answer. If what the computer produces is the **exact** solution of **just as good a model** of the physical system as was originally written down, we can get just as much insight from the computer solution as we can from the exact solution of the originally specified problem.

The idea of allowing the calculator or computer to *change the problem*, albeit not by very much, in some norm, is upsetting to many mathematicians, because one of the most powerful ideas in mathematics is that the irrelevant details of the physical context of a problem can be ignored. However, most people will have to deal eventually with the fact that mathematical problems encountered in science

and engineering are usually merely one representative out of an infinite class of mathematical models for the phenomenon in question, and further that the input data to the model will usually be of low accuracy compared to the precision available on most computers or calculators. In such cases, fanatical obsession with accurately solving the specified model problem is neither necessary nor appropriate, while analysis of the effect of perturbations of the input data and/or the model is essential.

I remark that sometimes the method or the program or the computer will change the problem by a *large* amount, or by a small amount in a nonphysical way. For example, the computer may not preserve the physical invariants of the problem. In such cases, we say that the numerical method is at fault and must search for a better method.

NIEVERGELT'S PROBLEM REVISITED. The object of backward error analysis is to put computational errors into the context of modelling errors. However, no “real-world” context for Nievergelt’s example was given in [1], so what follows is speculative.

The elements of A are integers, and A is symmetric. Perhaps, then, A arose as the normal equations from some least-squares problem where the entries of the (possibly rectangular) matrix B where $A = B^T B$ were also integers. We found that if

$$B = \begin{bmatrix} 666 & 665 \\ 667 & 666 \end{bmatrix}$$

then $A = B^T B$. There may well be other such B (of different dimensions), as an exhaustive search was not carried out. (The occurrence of the digit “6” in the entries of B is one reason for the “Six” in the title, by the way.)

It is well-known in numerical analysis circles that using the normal equations can increase the condition number drastically, making the problem much more sensitive than it need be, and it is certainly the case here that A is much more ill-conditioned than B .

On the other hand, $\det(B) = \det(A) = 1$ and the entries are integers, so it may be that the problem was combinatorial in origin. In this case, nothing short of exact arithmetic is sufficient since *no* perturbation in the data is allowed. The fact that Cramer’s rule got the right answer for this problem in [1] was accidental: Cramer’s rule is unstable, even for the 2 by 2 case [6]. We note that there are many different ways of doing exact computations on computers nowadays, including the use of symbolic manipulation programs such as Maple, or perhaps interval arithmetic.

We now look a bit more closely at the apparently random digits of the answers produced by the various methods tried in the Nievergelt paper. Note that after the results of the different methods were given, the question was asked:

“What might instructors and textbooks tell students, who carefully copy down their supercalculator’s results without ever suspecting that they are copying random digits?”

If we compute the ratio x/y for each of the solutions produced, **including the exact solution**, we get $x/y = -0.9984996258$, *for all solutions except the singular value decomposition solution*. The solutions are *not* random digits! What, then, are they? Clearly the vector $[x \ y]$ produced is to high accuracy a multiple of some unit vector. What unit vector? The answer lies in a detailed look at the singular value decomposition. This might be beyond what is desired to teach, at least in an introductory course (see [7] for some details of the use of the HP28S at Western).

However, it is instructive to pursue this here. We first note that if the singular value decomposition (see [8]) of B is $B = U\Sigma V^T$ where U and V are orthogonal and $\Sigma = \text{diag}(\sigma_1, \sigma_2)$ where $\sigma_1 > \sigma_2 > 0$, then $A = V\Sigma^2 V^T$ and the singular values of A are the squares of the singular values of B . We can write the solution to $A\mathbf{x} = \mathbf{b}$ in terms of these singular values as

$$\mathbf{x} = \sigma_1^{-2} v_{11} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} + \sigma_2^{-2} v_{12} \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix}$$

and we note that $\sigma_1 \approx 1332.00075075033$ while $\sigma_2 \approx 0.00075075033$. The singular values were computed by MATLAB [9], and separately checked by hand and by Maple [5].

For small σ_2 , the solution is clearly dominated by the second term, and hence the near-constant ratio of the computed solutions observed earlier. However, if σ_2 is small compared to σ_1 , then the matrix is ill-conditioned, *and perturbations in the data will change the value of σ_2 drastically*. The singular-value decomposition solution obtained in [1] is just the first term above, on the other hand, and it is robust under perturbations.

It is remarkable that the solution produced by the HP keys directly correspond to the above with $\sigma_2 \approx 1/1601$, while using Gaussian elimination as programmed gives $\sigma_2 \approx 1/1413$, and the eigenvalue approach gives $\sigma_2 \approx 1/1332.13$, whereas the exact solution has $\sigma_2 \approx 1/1332.00075075$.

MORAL 1: Each method gives the exact solution for a slightly perturbed matrix, all with small σ_2 . Each solution is *just as good* as the exact solution, in view of the possibility of data error. Note particularly that this problem does *not* go away if you use another software package, *even if* it gives the exact solution.

MORAL 2: Explaining the unexpected behaviour of the calculator involves more serious mathematics than (perhaps) was anticipated. This seems to be one of the better arguments *for* using such equipment in a classroom setting, in that using the calculator can enrich the mathematical content of the course.

CONDITIONING OF COMPUTATIONAL PROBLEMS. It turns out that this behaviour of computational equipment is quite general. The idea of condition number can be used in any computational problem, to assess (to first order) the effects of perturbations in the model or the data. This is a good idea, even if you have the exact solution to your problem, and provides an excellent motivation to study perturbation theory. I give some examples below.

Function Evaluation. In evaluating the function $y = f(x)$, we consider the relative effect of a small change Δx on y . Of course this leads to the first order expression

$$\frac{\Delta y}{y} \approx \left(\frac{x f'(x)}{f(x)} \right) \frac{\Delta x}{x}$$

and the expression in brackets is called the condition number of f , and often denoted by C (see e.g. [3]). Note that this is just an application of the theory of differentials, a standard topic in most calculus courses.

This number C is an appropriate number to look at if x is not zero and $f(x)$ is not zero, in which case absolute errors, as opposed to relative errors, are the quantities of interest. Here, we look at only two examples, the tangent function, for which $C = x/(\sin(x)\cos(x))$, and the exponential function, for which $C = x$. Consider the exponential function first. If $x \approx 100$, then $C \approx 100$. If we know x only

to two figures, how accurately can we know $y = \exp(x)$? The relative error in y is on the order of 1. That is, we are not sure of any figures of our answer. Note that the precision used to calculate $\exp(x)$ is almost irrelevant, and in particular going to “double precision” doesn’t help. The difficulty is that *data error* is amplified, not that computational errors are amplified.

Now consider $\pi/2$ which is 1.57079632680 to the precision of the calculator. The calculated value of $\tan(\pi/2)$, using the calculator in radians, is -195948537906 . This result shows that someone took fanatical care with the programming of the calculator, because it is correct to all figures *given the assumption that we wanted the tangent of 1.570796326800000000... and not $\tan(\pi/2)$* . In degree mode, taking $\tan(90)$ gives TAN Error: Infinite Result, with the infinite result flag set. What is the condition number of $\tan(x)$ near $x = \pi/2$?

$$C = \frac{-\pi}{2(x - \pi/2)} - 1 - \frac{1}{3}(x - \pi/2) + O((x - \pi/2)^2),$$

giving clear evidence of the difficulty. In the light of this extreme sensitivity to data error, what purpose is served by getting 12 figures correct for $\tan(1.570796326800000000...)$? The assumption that all the unknown figures are zero is simply not always tenable.

In some sense, the paper [1] contended that the HP28S is not precise enough. The above example shows that in fact, it is too precise, at least for some problems.

Rootfinding. Wilkinson [10] has given a superbly clear account of the sensitivity of some polynomial roots to very small changes in the polynomial coefficients. I will not repeat the analysis here of the now-famous Wilkinson polynomial $p(x) = (x - 1)(x - 2) \cdots (x - 20)$, except to note that many people take the sensitivity of an input polynomial root as a motivation to use higher precision in rootfinding, perhaps even using arbitrary precision. The sad truth is that if the roots are ill-conditioned, then *small changes in the input data* can drastically change the roots, *irrespective of the solution method used*.

Solution of Differential Equations. As a final example, we look at some differential equations. These, too, have condition numbers, and not surprisingly we find that some ill-conditioned d.e.’s are of great interest: all chaotic problems are (by definition) exponentially ill-conditioned as initial-value problems. It turns out that a system is chaotic if the trajectories are bounded but $C \sim \exp(\lambda t)$ for some $\lambda > 0$. In fact, λ is precisely the largest Lyapunov exponent. This means that for a chaotic problem, data errors (i.e. in the initial condition), modelling errors, and computational errors are amplified at an exponential rate. This implies that a good approximate solution, good in the sense of being close to the “exact” solution of the specified problem, is impractical to compute, since good accuracy at the end of the range of integration requires exponentially high precision for the initial conditions (regardless of the solution technique—this would hold even if you knew the exact solution). However, in the context of modelling errors, and given the point of view that a good numerical method will give you the *exact* solution of a nearby differential equation (i.e. just as good a model of the underlying physical problem as the specified model), much insight can still be gained from such numerical computations, even though the computed solutions are practically guaranteed to be nothing like the exact solutions. See [11] for details.

So what does this all have to do with the calculator? This material is almost certainly too advanced for a first course in calculus, after all. However, there is an example of a differential equation that does come up in any first year course, and the HP28S even has a key for numerically solving it. Of course, this is the d.e. $y'(t) = f(t)$ —that is, compute the definite integral $\int_{t_0}^t f(\tau) d\tau$. The performance of this key is quite remarkable, and if the HP28S is used in class, the students like it very much indeed. The following example (perhaps even more horrifying than the linear algebra example of [1] at first glance) provides an excellent pedagogical opportunity.

Problem: compute, on the calculator, $\int_0^1 (d\tau/\tau)$.

The calculator will take more than an hour, almost regardless of the input error tolerance, to return its answer 22.82...—which, of course, is nothing like the correct answer of infinity.

This example is accessible to students very early on in the curriculum. Many are extremely startled by the peculiar behavior of the calculator. Of course, the calculator returns an error message (albeit in the very cryptic and easily overlooked form of a negative reported estimated error), so it is not quite as bad as it seems. However, this example provokes a very desirable degree of skepticism in the student, and gives strong motivation for the study of improper integrals (and analytic integration). The observation that the calculator is in fact giving them the *exact* answer to $\int_\epsilon^1 (d\tau/\tau) + \int_0^1 \delta(\tau)$ for $\epsilon \approx 10^{-10}$ and $\delta(\tau) \approx 10^{-12}$ provides them with welcome relief and understanding of what the calculator has done. After the relief comes the realization that most of the area of the figure lies next to the singularity, which is a valuable pedagogical point.

Conclusions. The point of view of backward error analysis, *i.e.* that a good numerical method gives the exact solution to a nearby problem, is very helpful in explaining the unexpected behaviour of calculators and computers on sensitive problems. This point of view is not a panacea, as some problem contexts preclude the necessary changes in the problem. In particular, Professor W. Kahan cautions that backward error analysis is intended only as *explanation* and not as justification, and warns that there are problems for which backward error analysis fails. A simple example is composition of functions—if we have a way of computing $f(x)$ with good backward error, and a way of computing $g(u)$ with good backward error, then it is not necessarily true that we can compute $(g \circ f)(x)$ with good backward error. But, there is no question that backward error analysis does help with a large class of practical problems.

This paper echoes the call of [1] for the teaching of the theory of conditioning of problems. Since this is merely an application of differentials, or the first term in a Taylor series, or the first term in a perturbation expansion, this task is actually desirable for several reasons.

In essence, this paper has shown that the difficulties exhibited in [1] were not the fault of the calculator, but rather the fault of the problem, in some sense. Further, these difficulties actually provide motivation for the student to learn sensitivity analysis and the use of differentials, norms, perturbation series, and other more sophisticated mathematical topics than those just to “solve” the problem. In view of data error, *these topics should be learned anyway*. It has also been made clear that the theory of condition numbers is not restricted to numerical linear algebra, but is in fact of wide practical utility. This usefulness is only enhanced by the existence of supercalculators, by which many of our students

are exposed to very powerful and sophisticated environments for scientific computation.

ACKNOWLEDGMENTS. I would like to thank Ms. Amanda Connell for programming assistance in finding *B*. Figure 1 was prepared by Professor George Corliss. Doug Moseley and Professor Chris Essex together suggested the title.

REFERENCES

1. Y. Nievergelt, *Numerical Linear Algebra on the HP-28 or How to Lie with Supercalculators*, this MONTHLY, vol 98, no. 6, pp. 539–543, June–July 1991.
2. D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software*, Prentice-Hall, 1989.
3. G. Dahlquist and Åke Björk, *Numerical Methods*, Prentice-Hall, 1974.
4. U. Ascher, R. M. M. Mattheij, and R. D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall 1988.
5. B. Char, K. O. Geddes, G. H. Gonnet, M. B. Monagan, and Stephen Watt, *The Maple Reference Manual*, 5th ed., WATCOM 1988.
6. G. W. Stewart, “Cramer’s Rule,” USENET posting to sci.math.num-analysis, No. 249, Message-ID: <42829@mimsy.umd.edu>, 12 Nov 91.
7. R. M. Corless, C. Essex, P. J. Sullivan, and P. A. Rosati, Use of the HP28S Supercalculator in First Year Engineering Mathematics Courses, *Proc. 7th Canadian Conference on Engineering Education*, Toronto, June 1990.
8. G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins, 1983.
9. The MATLAB Reference Guide, The MathWorks, 1989.
10. J. H. Wilkinson, *Rounding Errors in Algebraic Processes*, Prentice-Hall, 1963.
11. R. M. Corless, “Defect-Controlled Numerical Methods and Shadowing for Chaotic Differential Equations”, *Physica D*, vol. 60, 1992 pp. 323–334.

Department of Applied Mathematics
University of Western Ontario
London, Canada N6A 5B9

According to cable reports from London, the Council of Trinity College, Cambridge, has removed Professor BERTRAND RUSSELL from his lectureship in logic and principles of mathematics on account of his having been convicted under the defense of the realm act for publishing a leaflet defending the “Conscientious Objector” to service in the British army. Professor Russell is well known in this country through his mathematical writings.

—*American Mathematical Monthly*
23, (1916) p. 317

What Is a Napierian Logarithm?

Raymond Ayoub

§1. INTRODUCTION. The invention of logarithms in 1614 by John Napier, baron of Merchiston in Scotland, is one of those rare parthogenic events in the history of science—there seemed to be no visible developments which foreshadowed its creation. The subsequent progress completely revolutionized arithmetic calculations in various areas of science, especially in astronomy. It is startling to realize that the spectacular, if not miraculous, development of computers in the last two decades has rendered tables of logarithms, and the portable version—the slide rule—essentially obsolete.

This was not always so. A generation ago, the use of tables of logarithms was an integral part of secondary education. A student had to learn the meaning of the terms logarithm, base, antilogarithm, mantissa, and interpolation and had to learn to use the tables. Moreover, the tables were either to base 10, earlier called “Briggsian” or to base e earlier called “hyperbolic” and before that “Napierian”. The logarithms invented by Napier were closely allied to, but not the same as, hyperbolic.

Over the years, various authors have vied with one another to produce tables of greater precision as well as ease of use. Indeed as recently as 1964, a table of logs to 110 decimal places was published under the auspices of the Royal Society.

The purpose of this essay is to explain Napier’s discovery and in the process answer the question of the title. The writer is motivated in part by the fact that historical accounts are either sketchy or inaccurate or both. We shall refer to some of these at the appropriate place in the narrative. Napier’s ideas are, as we shall see, quite subtle and many writers have failed to appreciate their brilliance and depth.

Napier was born at Merchiston near Edinburgh in 1550 and died in 1617. It is worth noting that Descartes lived from 1596 to 1650 while Newton lived from 1642 to 1727, the *Principia* having been published in 1687. Thus the mathematical tools available to Napier were decidedly limited. We should stipulate that the laws of exponents were, by then, well understood. Napier’s ideas exhibited a remarkably clear conception of the logarithmic function, the term “logarithm” having been coined by Napier himself. This was at a time when the concept of function was only vaguely understood by the scientific community of his day.

Moreover, he perfected the notation for the decimal representation of numbers, his notation being essentially that in use today. The decimal representation of numbers had been earlier described by S. Stevin, who built upon earlier work, and whose notation was not as elegant as Napier’s.

Two books were published on logarithms. The first in 1614 was titled “MIRIFICI LOGARITHMORUM CANONIS DESCRIPTIO” which has been translated as “Description of the wonderful canon of logarithms.” This contains a table of

logarithms together with rules for the solution of triangles, both plane and spherical, with the use of the “canon.”

The second was published posthumously in 1619 and was titled “MIRIFICI LOGARITHMORUM CANNONIS CONSTRUCTIO” or “Construction of the wonderful canon of logarithms.” It is in this work that he gives an account of the method by which the table was constructed as well as the properties of his logarithmic function, properties essential to the construction. It is this account that we propose to analyse and upon which we shall elaborate.

Before proceeding, we should add parenthetically that Napier was well-known and highly esteemed in theological circles for his analysis and interpretation of the Book of the Revelation of St. John the Divine!

§2. THE PROBLEM. The end of the 16th and beginning of the 17th century was a period of profound astronomical research with such celebrated scholars as J. Kepler (1571–1630), Tycho Brahe (1546–1601), Galileo Galilei (1564–1642). The need for carrying out elaborate calculations involving trigonometric functions was very pressing. It was therefore urgent that some procedure be sought to shorten the labor required to perform these calculations. One such aid was the use of the identity $2 \sin A \sin B = \cos(A - B) - \cos(A + B)$ which was given the tongue-twisting name of prosthaphaeresis. There is some evidence that the method was used by Brahe and his assistant Wittich to whom the method is sometimes attributed. Another was the use of identity $4 AB = (A + B)^2 - (A - B)^2$, which is sometimes referred to as the method of “quarter squares”.

Clearly these were inadequate. Ideally, what was needed was a function from (R_+^*, \cdot) to $(R, +)$ which converted multiplication to addition, in other words, a logarithmic function.

With the laws of exponents in mind, the most obvious approach to defining such a function is to begin with a fixed real number c (which, in what follows, we take to be < 1), calculate $b_n = c^n$ ($n = 1, 2, \dots$) and call n the logarithm of b_n . Moreover we need only calculate c^n for $n = 1, 2, \dots, k$ where k is that value which makes c^k about $1/2$. For if $b = c^m > 1/2$, we find l so that $\alpha = b/2^l$ with $\alpha \leq 1/2$ and then $\log b$ is determined by $\log \alpha$ and $\log 1/2$.

Reasonable though this approach may be, it is subject to difficulties which are not easily overcome and which we now describe. Moreover, we begin by imposing reasonable restrictions.

- (i) The value of c should be chosen so that the calculation of c^n ($n \in \mathbb{N}$) is arithmetically simple.
- (ii) The value of c^n should not decay too rapidly i.e. the values c^n and c^{n+1} should be relatively close to one another.
- (iii) Given two values c^r and c^s , if a is such that, $c^r < a < c^s$, we require a method for finding α such that $a = c^\alpha$. The method should be accurate and easy to use.
- (iv) The labor of carrying out the needed computations should be within manageable bounds.

Napier defines a mapping, which we describe in detail below, and is thereby led to choose $c = 1 - \lambda$ with $\lambda = 10^{-\sigma}$ and $\sigma \in \mathbb{N}$. Let us stress at once that these powers are reference points and are the basis upon which the canon is constructed. As we shall see, the choice of σ determines the degree of accuracy of the final product.

If we write $a_m = \lambda^{-1}(1 - \lambda)^m$ ($m \in \mathbb{N}$), then a_m conforms to requirements (i) and (ii) for $a_m = a_{m-1} - \lambda a_{m-1}$. Since $a_{m-1}\lambda$ is merely a shift of decimal, the value of a_m is easily obtained from a_{m-1} by a simple subtraction. The arithmetic could hardly be simpler. Moreover, with a proper choice of σ , a_m can be made to decay as slowly as we please.

Let $b_n = c^n$, and suppose we take $\log b_n = n$. If then a is such that $c^{n+1} < a < c^n$, and we find $\log a$ by linear interpolation, the error E satisfies

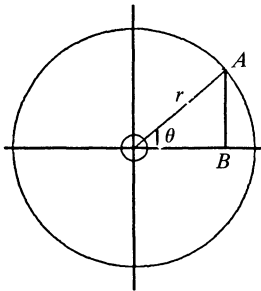
$$|E| \leq \frac{1}{2} 10^{-5} \text{ and is relatively sharp.}$$

Suppose c^k is about $1/2$ with $c = \lambda(1 - 10^{-5})$, then k is about 7×10^5 . If therefore we aim for 7 figures of accuracy, we must choose $\sigma = 7$ thus necessitating initially 7,000,000 calculations. This is an overpowering amount of labor. In addition, we have the labor of interpolation.

We state at once that this is not the method adopted by Napier.

It seems to this writer virtually certain that as Napier embarked on his project, he soon perceived the shortcomings of such a comparatively straightforward approach. He therefore sought and discovered an alternate route, deeper, more effective and ultimately successful. There is evidence that Napier started to think about the problem around 1594 but there is no record of any of his false starts. We should remark that a Swiss contemporary named JOST BURGI used this straightforward approach and published a table in 1620 well after Napier's work had been recognized and widely appreciated. Burgi's table proved to be of very limited use.

§3. PRELIMINARY REMARKS. Napier's motivation was the simplification of calculations related to the solutions of triangles, especially spherical triangles which were crucial in astronomy. In Napier's day and indeed for some time thereafter, the sine of an angle was not viewed as a ratio. It was taken to be the leg of a right triangle. More specifically, suppose we have a circle of radius r and an angle θ .



Then sine of θ was taken to be AB . The fact that the sine changed with the radius was not a serious impediment.

Napier sets out to construct a table which consists of the logarithms of $\sin \theta$. In this table Napier chooses $r = 10^7$ and his table consists of logarithm $10^7 \sin \theta$ ($30^\circ \leq \theta \leq 90^\circ$) and at increments of $1'$. Thus, although he has in mind applications to trigonometry, the table is in reality a table of logarithms ranging from 10^7 to $10^7/2$ with increments which are not arithmetic but "geometric". Finally we

note that although the words “cosine” and “tangent” had not yet come into use, his table includes logs of cosines and tangents.

Moreover, let us stress that we have used the word “logarithm” generically. Napier’s logarithms, which we shall denote below by $LN(x)$, have somewhat different properties from the standard function $\log x$, which as we know, defines an isomorphism of (R_+^*, \cdot) and $(R, +)$. These differences are not significant but have led to misinterpretations by some historians. Even the redoubtable French general has implied mistakenly, that Napier’s function was an isomorphism from (R_+^*, \cdot) to $(R, +)$. This assumption is also inherent in other authors’ assertion that Napier chose the base e or the base $1/e$. Though not an isomorphism, the fundamental idea, however, of accurately converting multiplication to addition is essentially preserved.

§4. THE CONSTRUCTION. To determine a correspondence between a set in “geometric progression” and one in “arithmetic progression”, that is, an exponential mapping which is the key element in defining a logarithmic function, Napier ingeniously resorts to a model from mechanics. It is based on the simple idea that the displacement of a point which moves with constant velocity is “arithmetic” while the displacement of a point which moves with a velocity proportional to the displacement is “geometric”. The correspondence between two such points defines the required mapping. Here is Napier’s model. Let TS be a segment of fixed length

$$w = 10^7.$$

The choice of the fixed length TS is motivated by the fact that Napier was interested in logarithms of $\sin \theta$ and is not a whimsical one! Given his objective, the choice was perfectly reasonable, the value for w being dictated by the degree of precision desired.

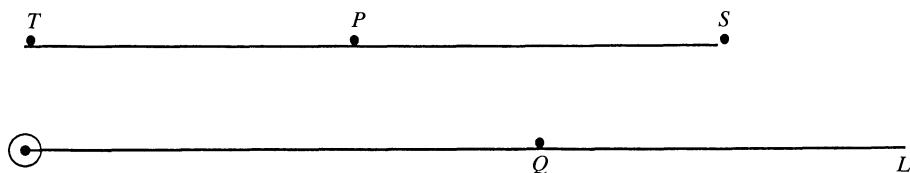


Figure 1

Let OL be another line extending to infinity. At $t = 0$, let the points start at T and O respectively. The point Q moves on OL with constant velocity v_0 . The point P moves on TS in such a way that its velocity is proportional to the distance PS and we assume its velocity at T to be v_0 . The velocity of P therefore, decreases from v_0 at T to 0 at S .

If at a time t , the point P is at a distance x from S , i.e. $PS = x$ and Q at a distance y from O , i.e. $OQ = y$, then Napier defines

$$y = \text{logarithm of } x.$$

We shall write

$$y = LN(x) \tag{1}$$

to distinguish this function from the standard logarithm. We shall shortly establish the connection between the two.

Evidently

$$LN(w) = 0, \quad (2)$$

since when P is at T , we have $PS = w$ while $OQ = 0$.

Before proceeding to the table of logarithms, let us use mathematics not then available to Napier to see exactly what his function $LN(x)$ is.

We have by Napier's conditions,

$$\frac{dx}{dt} = -kx$$

and at $t = 0$, $x = w$ and $dx/dt = v_0$. Hence $k = v_0/w$.

Since

$$\frac{dy}{dt} = v_0, \quad \text{we get}$$

$$\frac{dy}{dx} = -\frac{w}{x} \quad \text{or} \quad y = -w \ln x + c$$

But when $t = 0$, $x = w$ and $y = 0$, hence $c = w \ln w$ or finally

$$\begin{aligned} LN(x) &= w(\ln w - \ln x) \\ &= \ln \left(\frac{w}{x} \right)^w. \end{aligned}$$

Many writers correctly state this fact about Napier's logarithms but it is not very illuminating except to verify that his logarithmic function is not an isomorphism. It is moreover, disingenuous to imply that he clumsily used a complicated function when he could have used a simple one! From (6) we have

$$LN(ab) = LN(a) + LN(b) - w \ln w;$$

hence

$$LN(ab) \neq LN(a) + LN(b).$$

A simple transformation however, gives $LN(x)$ a more familiar form. Let $\bar{y} = y/w$, $\bar{x} = x/w$, then, from (6)

$$\bar{y} = -\ln \bar{x} = \log_{1/e} \bar{x}.$$

This fact has led many writers to state erroneously, that Napier chose e or $1/e$ as the base of his logarithms. Napier's function is not an isomorphism, though, as we have remarked above, closely related.

It is interesting to observe that a modification of Napier's model will lead easily to the natural logarithm. Namely, assume that the velocity of P is proportional to OP while that of Q is constant. Assume too, that at $t = 0$, $OP = 1$ and $Q = 0$. The underlying differential equation has the natural logarithm as its solution. The reader may assign this as an exercise to a class in calculus.

§5. PROPERTIES OF $LN(X)$. Napier begins by doing some geometry in order to justify the claim that as P moves geometrically, Q moves arithmetically. We interpret his reasoning as follows.

Let P_1, P_2, P_3 be points on TS in geometric progression i.e.

$$P_1S : P_2S = P_2S : P_3S \quad (1)$$

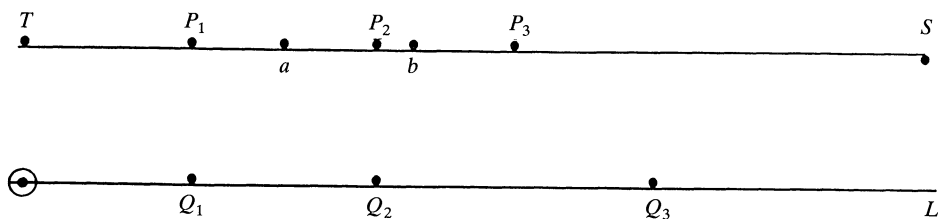


Figure 2

Let the corresponding points on OL be Q_1, Q_2, Q_3 .

Let a be an arbitrary point in P_1P_2 and b the corresponding point in P_2P_3 , that is

$$aP_1 : aP_2 = bP_2 : bP_3 \quad (2)$$

Let v_x be the velocity of the point P when it is at x ($= PS$). From (1) we get

$$P_1P_2 : P_1S = P_2P_3 : P_2S, \quad (3)$$

that is,

$$P_1S : P_2S = P_1P_2 : P_2P_3 = \lambda \text{ (say)}. \quad (4)$$

On the other hand, from (2) we have

$$P_1P_2 : aP_1 = P_1P_2 : P_2P_3 = \lambda \quad (5)$$

By construction, and from (4) and (5)

$$\begin{aligned} v_a : v_b = aS : bS &= P_1S - aP_1 : P_2S - bP_2 \\ &= \lambda P_2S - \lambda bP_2 : P_2S - bP_2 = \lambda. \end{aligned}$$

Since $P_1P_2 : P_2P_3 = \lambda$, it is very plausible to assume, and indeed Napier does assume, that the point a traverses the segment P_1P_2 in the same time that b traverses P_2P_3 .

Since the velocity on OL is constant, and we have shown that the times to traverse P_1P_2 and P_2P_3 are the same, it follows that

$$OQ_2 - OQ_1 = OQ_3 - OQ_2.$$

This conclusion forms the basis for Napier's further developments. The reader will note that Napier has, in effect, integrated the underlying differential equation.

We use the above analysis to derive properties of $LN(x)$, properties essential in the construction of the table of logarithms.

Theorem 5.1. *If $x_1x_4 = x_2x_3$, then*

$$LN(x_1) + LN(x_4) = LN(x_2) + LN(x_3) \quad (6)$$

Proof: Referring to figure 2, let

$$P_1S = x_1, P_2S = x_2, P_3S = x_3 \text{ with } x_1 : x_2 = x_2 : x_3,$$

and let $OQ_1 = y_1, OQ_2 = y_2, OQ_3 = y_3$, then

$$y_i = LN(x_i) \quad (i = 1, 2, 3),$$

and since $y_2 - y_1 = y_3 - y_2$, we get

$$LN(x_2) - LN(x_1) = LN(x_3) - LN(x_2). \quad (7)$$

Now choose x_4 so that

$$x_2 : x_3 = x_3 : x_4,$$

then $x_1 : x_2 = x_3 : x_4$ and

$$LN(x_3) - LN(x_2) = LN(x_4) - LN(x_3). \quad (8)$$

Combining (7) and (8) we get the result. Dropping the subscripts, we have that if $ab = cd$, then $LN(a) + LN(b) = LN(c) + LN(d)$.

Although $LN(x)$ does not satisfy the additive condition, it does however, satisfy the following modified property:

Theorem 5.2. *If TS is denoted by w , then*

$$LN(wab) = LN(wa) + LN(b)$$

Proof: Using theorem 1, we have

$$LN(wab) + LN(1) = LN(wa) + LN(b)$$

and

$$LN(wb) + LN(1) = LN(w) + LN(b)$$

Since $LN(w) = 0$ the result follows.

Cor.

$$LN(wc^n) = nLN(wc)$$

for $n = 0, 1, 2, \dots$

The proof is a straightforward induction.

The next step in the construction is to find bounds for $LN(x)$. These bounds are absolutely crucial to the calculation of the table of logarithms.

Theorem 5.3. *We have the following inequalities:*

$$x \left(\frac{w}{x} - 1 \right) < LN(x) < w \left(\frac{w}{x} - 1 \right) \quad (9)$$

Proof: Referring to fig. 3, if P and Q are corresponding points, then because P is slowing down from its initial velocity v_0 at T , and Q is moving with constant velocity v_0 , it follows that $OQ > TP$. Hence

$$y = LN(x) > TP = TS - PS = w - x.$$

On the other hand imagine that the point P goes to P' , a point to the left of T and let Q' be the corresponding point on OL , then the velocity at P' is greater than v_0 , and hence

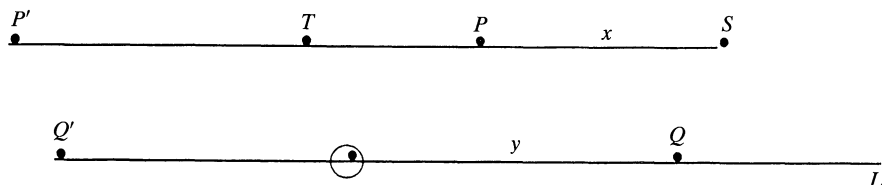


Figure 3

$$OQ' < TP'.$$

But from the geometry, we know that if $OQ' > OQ$, then

$$TS: PS = TP': TP,$$

and hence

$$TP' = \frac{w}{x}(w - x).$$

Consequently

$$LN(x) = OQ = Q'O < TP' = \frac{w}{x}(w - x)$$

as required.

Cor. If $a < b$, then

$$\frac{w}{b}(b - a) < LN(a) - LN(b) < \frac{w}{a}(b - a). \quad (10)$$

Proof: Choose c so that $bc = aw$, then

$$LN(b) + LN(c) = LN(a).$$

Now apply the theorem to $c = aw/b$ to get the stated conclusion.

The reader will note once again Napier's remarkable insight since (10) gives the inequalities which follow from the mean-value theorem for $LN(x)$.

§6. THE CANON. The construction of the table takes place in four stages. The first is the calculation of a number of reference points. The second is the evaluation of the logarithms of these reference points. The third is the calculation of logarithms of intermediate values and finally the fourth step is the determination of logarithms lying outside the table.

Step 1. The reference points. These are the points

$$w(1 - c)^n$$

To understand the choice of these points, it suffices to take short time intervals and assume that the velocity of P in fig. 1 is constant in that interval. If Q_n is the position of the point Q which is n units from 0, and P_n the corresponding position of P , then it is readily seen that an approximation to P_n is given by $w(1 - c)^n$. Thus while $LN(P_n) = n$, we stress emphatically that $LN(w(1 - c)^n) \neq n$. The assertion that $LN(w(1 - c)^n) = n$ is one of the most flagrant errors made by commentators.

We have already observed that the calculation of the values

$$A_n = w(1 - 10^{-k})^n$$

for a fixed k is arithmetically comparatively simple and this fact reinforces Napier's decision to choose these points.

We have seen that in order for $w(1 - c)^n$ to be approximately $w/2$, n must be about 7,000,000 (with $c = 10^{-7}$). This is an overwhelming amount of calculation. Since, as we now know,

$$LN'(x) = -\frac{w}{x}, \quad (1)$$

and in the range $1/2 \leq x \leq 1$, $LN'(x)$ varies from $= w$ to $-2w$, Napier feels justified in taking larger "bites" in getting to $w/2$. This tactic, explained below,

enables him to get to $w/2$ in 1600 steps and, as it turns out, retain seven decimal places of accuracy for these reference points.

Here then are the steps taken by Napier. He calculates 3 tables as follows.

Table I. Calculate $a_n = w(1 - c)^n$ for $n = 0, 1, 2, \dots, 100$. The last entry in this table is $w(1 - 10^{-7})^{100}$ and this is approximately $w(1 - 10^{-5})$. The reader should recall that if a_n has been calculated then $a_{n+1} = a_n - 10^{-7}a_n$. The second term is merely a shift of decimal. The same remark applies to the remaining tables.

Table II. Calculate $b_n = w(1 - 1/10^5)^n$ ($n = 0, \dots, 50$). The last entry in this table is approximately $w(1 - 1/2000)$.

Despite the simplicity of calculating the values b_n , Napier makes an arithmetic error which affects the accuracy of subsequent calculations. We shall not dwell on this point since it has no bearing on the validity of his method.

Table III. This is a double array

$$c_{m,n} = w \left(1 - \frac{1}{2000}\right)^{m-1} \left(1 - \frac{1}{100}\right)^{n-1}$$

for $1 \leq m \leq 21, 1 \leq n \leq 69$.

The array has 69 columns and 20 rows. The first column begins at approximately the point where Table II ended and since $(1 - 1/2000)^{20}$ is approximately $(1 - 1/100)$, the last entry of any column is approximately the second entry of the next column.

The last entry in the last column is

$$w \left(1 - \frac{1}{2000}\right)^{20} \left(1 - \frac{1}{100}\right)^{68},$$

and this is about

$$a = w \left(1 - \frac{1}{100}\right)^{69}$$

or about $w/2$.

The computation of the reference points of Tables I, II and III involves 1600 calculations, a far cry from 7,000,000.

Step 2. The Radical Table.

Having calculated the reference points, Napier now proceeds to construct what he calls the Radical Table, that is to say the logarithms of all the reference points of table III. He does this in a systematic way with the help of tables I and II.

First we calculate the logarithms of entries of tables I and II. Let us begin with table I. Using the inequalities of theorem 3, we find

$$1 < LN(w(1 - c)) < 1.0000001,$$

Napier takes the value to be the arithmetic mean i.e.

$$LN(w(1 - c)) = 1.00000005 \quad (2)$$

which is accurate to 14 decimal places.

If we systematically calculate $w(1 - c)^k$ ($k \leq 100$) recursively we find $w(1 - c)^{100} = 9999900.0004950$.

Using the Corollary of Theorem 5.2, we get $LN(w(1 - c)^k) = kLN(w(1 - c))$ and in particular

$$LN(w(1 - c)^{100}) = 100.000005 \quad (3)$$

We have lost 2 decimal places in the process. Thus we have all the values of Napier's logarithms for the points in table I.

Now we calculate the logarithm of the second entry of table II viz

$$b = w(1 - 10^{-5}) = 9999900.$$

This is based on the fact that b is approximately equal to

$$a = w(1 - c)^{100} = 9999900.0004950, \text{ whose logarithm we determined in (3).}$$

Writing

$$LN(b) = LN(a) + (LN(b) - LN(a)),$$

Napier estimates $LN(b) - LN(a)$ using the inequalities of Theorem 5.3. To 10 decimal places, we find

$$LN(b) - LN(a) = .0004950,$$

and therefore we evaluate

$$LN(b) = 100.0005.$$

The logarithms of all entries in table 2 are now obtained using the Corollary of Theorem 5.2, i.e. $LN(w(1 - 10^{-5})^k) = kLN(w(1 - 10^{-5}))$.

We pass to Table III. The last entry of Table II is

$$y = w(1 - 10^{-5})^{50} = 9995001.224804023027881. \quad (4)$$

The second entry of Table III (the first is 1) is

$$x = w\left(1 - \frac{1}{2000}\right) = 9995000 \quad (5)$$

which is approximately $w(1 - 10^{-5})^{50}$.

To find its log, we could proceed as above, but this would result in a significant loss of accuracy. So Napier introduces a subtle idea which we describe. Because

$$LN(x)' = -w/x,$$

the accuracy of the inequalities of Theorem 5.3 is much greater the closer x is to w . Napier uses this observation (which he has evidently discerned) as follows:

Assuming $LN(y)$ is known find the value of $LN(x)$. First choose z satisfying

$$\frac{w}{z} = \frac{y}{x}. \quad (6)$$

It is easily seen that z lies in the range of table I. Since from (3)

$$LN(z) = LN(x) - LN(y),$$

writing

$$LN(x) = LN(y) + LN(z),$$

reduces the computation to $LN(z)$. Let us call this the "method of transfer". To illustrate, let us show how to calculate $LN(x)$ of equation (5) given $LN(y)$ of equation (4). From (6), we have,

$$z = 9999998.77458344.$$

The value in Table I which is closest to z is $a = 9999999$ whose logarithm we found to be $LN(a) = 1.00000005$. Using the inequalities of Theorem 5.3 applied to a and z , we find

$$LN(x) = 5001.2504168229,$$

accurate to an astonishing 10 places of decimals.

Thus using the Corollary of Theorem 5.2, the logs of column 1 of Table III are immediately evaluated. We use the method of transfer on the last entry of column i to the second entry of column $i + 1$. The remaining logs are evaluated using the corollary of Theorem 5.2. These logarithms are accurate to 7 places of decimals.

Step 3. This step consists of interpolating at intervals of $1'$. This is done using the inequalities of Theorem 5.3. In addition there is a labor saving device using the identity

$$LN\left(\frac{w}{2}\right) + LN(w \sin 2\theta) = LN(w \sin \theta) + LN\left(w \sin\left(\frac{\pi}{2} - \theta\right)\right),$$

which enables us to read off the values for $30^\circ < \theta \leq 45^\circ$ from those already calculated.

In §7, we shall comment further on the accuracy of all the entries.

Step 4. The Short Table.

The final step is to find logarithms of numbers not in the range $[w/2, w]$. To do this Napier constructs what he calls the “short table”. It consists of the values

$$I(A) = -LN(Aa) + LN(a)$$

for

$$A = 2^p \times 10^q \quad 0 \leq p \leq 3, \quad 0 \leq q \leq 7.$$

$I(A)$ does not depend upon a as is easily seen.

Suppose for example that $0 < a < w/2$.

Choose m so that

$$\frac{w}{2} \leq 2^m a < w.$$

Then $2^m a$ lies in the range of the radical table, and

$$\begin{aligned} LN(a) &= LN(2^m a) - LN(2^m a) + LN(a) \\ &= LN(2^m a) + I(2^m). \end{aligned}$$

Thus knowing $I(2^m)$ permits us to evaluate $LN(a)$.

To find $I(A)$, we begin with

$$wa = \frac{w}{2}(2a).$$

Then by Theorem 5.1,

$$LN(a) = LN(2a) + LN\left(\frac{w}{2}\right)$$

(recall that $LN(w) = 0$).

From the radical table however, we find

$$LN\left(\frac{w}{2}\right) = 6931469.22.$$

This gives $I(2)$. By induction we get $I(2^k)$. From the relation $2^3aw = (8w/10)10a$, we find $LN(a) - LN(10a) = I(10) = 23,025,814$, and by induction $I(10^k)$, and finally $I(A)$.

A reader who has worked recently with logarithms, will not fail to recognize that

$$6,931,469 = 10^7 \ln 2$$

while

$$23,025,814 = 10^7 \ln 10.$$

Finally therefore, all information is available to calculate $LN(a)$ for any value of a .

Let us finally try to summarize Napier's method. Using a model based on mechanics, Napier defines a function from $A = [0, w]$ ($w = 10^7$), to \mathbb{R} . Identities satisfied by $LN(x)$ are proved and in effect the Mean Value Theorem for $LN(x)$ is used to derive inequalities satisfied by $LN(x)$. These are logarithmic-like properties.

To evaluate $LN(x)$ a subset $S \subset B = [w/2, w]$ consisting of about 1600 reference points is chosen. This subset is not random but is generated by powers of a number $c < 1$ so that control over the set is maintained. Using properties of $LN(x)$, the values $LN(x)$ for $x \in S$ are calculated. Using a subtle interpolation scheme, $LN(x)$ for $x \in B$ is calculated at intervals of 1'. Finally a table is given to facilitate the calculation of values of x which lie outside of B .

§7. BRIEF ERROR ANALYSIS. How accurate are Napier's logarithms? It is more pertinent to ask how accurate his method is for, as we have observed, he made an error in table II which affected subsequent calculations.

We begin by estimating the error of the value of $LN(x)$ assumed by Napier. Recall that $LN(x)$ is a decreasing function. Consider two values $b > a > 0$. To calculate $LN(b)$ given $LN(a)$, we write

$$LN(b) = LN(a) - D,$$

where

$$D = LN(a) - LN(b).$$

From Napier's inequalities of Theorem 5.2, we have

$$\frac{w}{b}(b-a) < D < \frac{w}{a}(b-a).$$

Napier takes the mean of the bounds

$$A = w/2(b-a)\left(\frac{1}{b} + \frac{1}{a}\right).$$

(In fact he often divides not by a or b but by some intermediate values to render the arithmetic easier.)

On the other hand, using Taylor's theorem we find that the difference is

$$A - D = E = \frac{w(b-a)}{2} \left(\frac{1}{a} + \frac{1}{b} \right) - \frac{w}{a}(b-a) + \frac{w(b-a)^2}{2a^2} - \frac{w(b-a)^3}{3a^3} + \dots$$

A little calculation shows that

$$E = \frac{w(b-a)^3}{6a^3} + \frac{w(b-a)^4}{a^3} \left(\frac{1}{4a} - \frac{1}{2b} \right) + 0 \left(\frac{(b-a)^5 w}{a^4} \right),$$

from which we may verify that the value

$$LN(9,999,999) = 1.00000005$$

is indeed accurate to 14 places of decimals. A simple further analysis confirms that the reference points are accurate to 7 places of decimals.

For the interpolated values, the error E is bounded by

$$|E| < 10^{-4}.$$

Thus the interpolated values are accurate to 4 places of decimals. However in the neighborhood of $a = 10^7 \sin \theta$ with θ close to 90° , the accuracy is far better—we saw that above.

We could improve the accuracy by the method of transfer which entails considerably more labor. Napier, however, is content to settle for the accuracy he has achieved.

In fact, if we use the method of transfer to calculate $LN(5,000,000)$, we find

$$LN\left(\frac{w}{2}\right) = LN(5,000,000) = LN(10^7 \sin 30^\circ) = 6931471.8055994$$

accurate to 7 decimal places. Note that $w/2$ is not one of the reference points.

The relative error of Napier's method is 10^{-10} .

§8. CONCLUDING REMARKS. Not surprisingly, the canon was greeted with great enthusiasm especially by Kepler who had been laboriously making calculations in connection with his laws.

It has the drawback we mentioned that $LN(1) \neq 0$. This drawback is an impediment to the ease of calculations but not a serious one. This weakness was recognized by Napier and during a visit by John Briggs to Napier, they discussed the possibility of constructing a table in which $\log 1 = 0$. This was subsequently completed by Briggs since Napier died soon after the visit.

The first table of hyperbolic logarithms, i.e. to base e , was first published by John Speidell in 1619. It was derived directly from Napier's table.

The odyssey of log tables is an interesting one but we shall not add any further details here.

While it is true that Napier's table was quickly overshadowed by others which were easier to use, it is well to bear in mind that it was Napier who, alone, led the way for others to follow. John Briggs' praise of Napier is one witness to this fact.

Finally, however, we cannot resist the temptation of quoting Napier's advice on forming a logarithmic table.

"Prepare forty-five pages, somewhat long in shape, so that besides the margins at the top and bottom, they may hold sixty lines of figures. Then divide each

page Next write on the first page at the top, to the left, over the first three columns, '0 degrees'; and at the bottom”.

Alas such charm has virtually vanished from the pages of our journals.

BIBLIOGRAPHY. The entire essay is based on “The Construction of the Wonderful Canon of logarithms and their relations to their own natural numbers.”

I have used the translation of William Rae MacDonald first published in 1889 and reprinted in 1966 for Dawson’s of Pall Mall, London. The translation also contains a catalogue of Napier’s works.

Since writing this article the author has found a reference to a lecture written by E. W. Hobson entitled “John Napier and the invention of logarithms, 1614. Cambridge Univ. Press 1914. It is to be highly recommended.

*Department of Mathematics
Pennsylvania State University
University Park, PA 16802*

The Mathematical Association of America wishes again to call the attention of all its members to the working arrangement between the Association and the *Annals of Mathematics* by which, in return for a certain subsidy contribution from the Association, the *Annals* has extended the size of its volume to include approximately one hundred pages of expository articles and at the same time has made the special subscription rate to individual members of the Association of one half the regular price. A goodly number of Association members have already taken advantage of this reduced rate, but it is felt that a much larger number would probably do so if their attention were sufficiently arrested.

—*American Mathematical Monthly*
32, (1925) pp. 324

The Tyranny of Tests

Peter Hilton

Disclaimer This article is full of generalizations, many about student attitudes. Such generalizations, unlike generalizations in mathematics, are not invalidated by the existence of counterexamples. All of us know the joy of teaching those exceptional undergraduates who really want to learn and are stimulated by the beauty and power of the mathematics they are learning. It is these students who give us the conviction that we are, as teachers, doing something thoroughly worthwhile, despite our many failures.

1. INTRODUCTION. I have been distressed for many years by the defects in our methods of evaluating students—at all levels from 3rd grade upwards to graduate students—and have railed especially against the unfortunate effects of over-frequent testing and inappropriate tests, especially multiple-choice and standardized tests. I have not changed my point of view on these issues; but I have come to realize that the ubiquity of intrusive testing, contaminating the learning environment, distorting the curriculum, and undermining the fabric of the teacher-student relationship, springs from a deep malaise in our whole approach to education, and is not to be remedied simply by attempting to improve the quality and relevance of the tests administered and reducing their frequency. Such a superficial approach may be thought of as treating a symptom and leaving the cause of the disease untouched. In this article I would like to develop this thought, and then raise the question of whether any change—any beneficial change, that is—is possible. I will largely confine myself to undergraduate education—principally, of course, in mathematics.

2. THE ROOT CAUSE OF THE PROBLEM. I have often remarked that, in this country, there is a fundamental confusion of education with training. Thus, whereas drivers of lethal automobiles stand in evident need of first-class training, we provide ‘driver education’; and whereas teachers stand in evident need of first-class education, we provide ‘teacher training.’ Indeed, I would now go further and say that, in general, our students—and their parents—only feel comfortable with the presence of a curricular item if it has an evident value for the students’ training.¹ It is plain why we would want our students to be able to use a word-processor, or to be familiar with the principles of business management, but of what *use* to them is the study of history or of literature?

Thus those subjects which are not obviously useful, in the sense of directly enhancing the student’s prospects of lucrative employment, are treated merely as hurdles on the path to success, not as sources of student enrichment. A prerequisite for success is, of course, the award of the degree. Thus the expectation of this award motivates the student and is exploited by the teacher as a spur to induce student effort. What has happened, then, is simply this—certification has replaced

¹So, apparently, do journalists if the jaundiced views of William Raspberry on the study of algebra are any guide.

education as the stimulus administered to the student, and hence as the perceived purpose of the learning experience. The student seeks the qualification of the degree and is not, generally speaking, overimpressed with the importance to him, or her, of the knowledge gained or the understanding achieved.

We, as mathematics teachers, support this system, since it would be unfair to our students to do otherwise. Neither in the questions we set in the tests we administer in our undergraduate courses, nor in our grading policy, do we insist that firm, usable knowledge be acquired. We give partial credit, knowing full well that the student could not have produced a complete solution. We set questions closely resembling problems discussed in class or assigned as homework. We rarely set a genuinely theoretical problem—certainly not an unfamiliar one. And, finally, we exploit a flexible grading policy allowing us to pass students who we know will never use any of the mathematical techniques they have ostensibly learnt—indeed, it is often a relief to us to feel this assurance. The notion that useful, usable knowledge has been transmitted from instructor to student scarcely enters into the final reckoning—but a qualification has been earned.

The future employer plays his part, too, in this distorted version of the true educational experience. He, or she, asks for transcripts of the candidate's performance as a student, and pays close attention to 'grade point average'. The actual content of the courses taken by the candidate, their real intrinsic difficulty, is rarely discussed. Moreover, this average is calculated on the basis of the candidate's performance over the four-year program for the bachelor's degree, so that improvement, resulting in successive grades of *C*, *B*, *A*, is no more advantageous than deterioration, giving rise to the sequence *A*, *B*, *C*. In particular, a student's mediocre performance as a freshman can never be expunged, and permanently contaminates the record.

Is it any wonder, then, that students are so desperately concerned about their grades, to the virtual exclusion of all other incentives to significant learning and good performance? We certainly cannot blame the students for endeavoring to maximize their expectations within the system as they find it; and we should remember that this is the same system, *mutatis mutandis*, as that with which they have become all too familiar during their pre-college education. We also cannot blame the students for choosing easy, undemanding courses in preference to difficult, challenging ones, where they have an option, since this is plainly their optimal strategy in the light of the perceived purpose of their undergraduate studies.

Thus the system ensures that students work for tests, and it is the conviction of most instructors that, in general, students will not work hard unless there is a test to prepare for. Certainly, there is a well-founded belief that students, having limited time to allocate to their studies, will surely favor those courses in which frequent tests are given. Thus the system has a built-in inertia militating towards its perpetuation.

There is a related issue concerning the administration of mathematical tests. Many instructors are obsessed by the necessity to ensure that students behave honestly. Because of this obsession they insist that tests take place under conditions under which no mathematician would ever contemplate doing mathematics. Moreover, university authorities usually mandate that the final examinations on any course are held at a fixed time in an assigned location under adequate invigilation. We prepare our students for this experience by insisting on the same conditions in the preliminary tests. An inevitable consequence of this is that we favor the student who thinks quickly over the student who thinks deeply, and we

reward accuracy far more highly, and more often, than we reward inventiveness and originality.

As I have already suggested, the pre-eminent importance of success in tests in the students' view of the purposes of education has a very deleterious effect on faculty-student relations, and perhaps especially in mathematics. Nobody who has been through the embarrassing experience of having to deal, following the announcement of examination results, with large numbers of students seeking to raise their scores by claiming partial credit for a piece of work which, while almost completely illegible, they maintain contains more grains of sense than the examiner allowed, can doubt this. Nor can any instructor who, having explained a particularly delicate, or particularly important, piece of theory, asks 'Are there any questions?' and receives the single question from many sources, 'Will we be asked that on the exam?'—nor can any instructor who has had that experience believe that his, or her, principal function is the transmission of knowledge and understanding; and that the student views the instructor as an ally in the educational process.

3. CAN WE IMPROVE THE SITUATION? It follows from my argument thus far that the problem is fundamentally a societal problem—there cannot be a deep, significant improvement until society better understands the purposes of education. We will make no progress at all if we merely set ourselves unrealistic goals and measure our success with defective instruments. Nor is it efficacious to couch our objectives in crude and absurd slogans. No 'education president' is going to alter the situation by declaring that by the year 2,000 we will be the best in the world in turning out well-educated science and mathematics students. Indeed, our only hope of achieving this objective lies in the visible decline of many of the nations with, hitherto, good educational standards and systems, as their citizens taste the fruits of capitalist democracy, and seek goals in life which education will not help them to achieve.

I do not know how to change our society in the fundamental way needed to solve the problem—it may be that it cannot be done. Thus, realistically, I can only hope to ameliorate the situation by proposing better tests, under more appropriate conditions, and occurring with less frequency. Indeed, I am most comfortable in proposing better tests at the pre-college level, since the grossest and most egregious errors in designing and administering tests of mathematical ability, knowledge and understanding occur there.

I am totally opposed to any test which gives the student no chance to explain how and why he or she arrived at the solution, or to make any other relevant comment on their answer. For it is precisely the student's process of thought that I wish to infer from the test. This is the human aspect of learning, as distinct from the machine aspect. So long as we are educating human beings and not machines, we should seek a measure of their capacity to apply their human intelligence. It is, of course, common ground that the virtual impossibility of devising multiple-choice standardized problems in geometry at the secondary level has played a decisive part in aborting the study of geometry in our secondary schools—to our immense loss.

However, the issue of multiple-choice tests also arises at the undergraduate level. This is due to the idealistic—some would say quixotic—principles which govern higher education in this country, and which decree that we should attempt to give post-secondary, *higher* education—not merely *further* education—to the majority of those graduating from high school. This is not the place to argue the

pros and cons of this educational and social philosophy; but it is appropriate to note the inevitable consequence that our universities and colleges have to deal with huge numbers of students at the freshman and sophomore level whose mathematical backgrounds are rather weak. The large numbers make it impractical—with the size of faculty which the current funding situation for higher education permits—to administer tests which require careful grading involving refined judgment. Thus machine-grading becomes involved, and thus the spectre of the multiple-choice test invades higher education. I have been much impressed by the ideas of Barry Cherkas and Joseph Roitberg at Hunter College, City University of New York, who have introduced an ingenious kind of multiple-choice question paper in which students can qualify for partial credit (see [CR]). This is a great gain. It increases student confidence and diminishes the risk of success through random guessing. It also goes some way to meeting my earlier objection to multiple-choice questions by allowing the examiner, rapidly and essentially algorithmically, to draw inferences about the students' thought-processes. It is plain from the students' own comments on this innovation that it is universally popular and strikes the students as more fair than either a traditional multiple-choice test or even a hand-graded test of the type familiar to us all. However, these student responses make it clear that they expect to get some credit even if they plainly have no hope of solving the given problem, and that they think it unfair that they should be expected to provide 'complete solutions' in any substantial numbers. Plainly, mathematics is not something they are learning how to *do*, or *understand*, but rather, as I have said, a means of ultimately achieving certification.

Even the ingenious innovation of Cherkas and Roitberg leaves another of the main objections to multiple-choice questions untouched. Very often, the optimal strategy for solving a mathematical problem when presented as a multiple-choice question is quite different from that of solving the same problem when presented without the artificial concomitant of a finite list of possible answers. One may simply test the possibilities given and either find one which definitely works, or eliminate all but one on the grounds that each is clearly wrong. An example of the former is to be asked to solve the equation $9/x = x/4$, where substitution of the hypothetical solutions is quick and effective; an example of the latter is to be asked to identify the curve which represents a given function $y = f(x)$, where one may show that all possibilities but one are inconsistent with certain obvious features of the function f .

Thus the difficulties created by the presence of very large numbers of students ensure that we can, as already mentioned, only hope to ameliorate the situation created by tests. However, as I have also argued, this is in any case the best we can hope to do in view of the central role played by tests, and testing, in our educational curriculum. I often find myself thinking that many of my colleagues seem more concerned with being able to demonstrate objectively that the students haven't learnt something than with successfully teaching it. I regard the imparting of knowledge and understanding as very much more important than devising means of determining whether or not I have done so successfully. I further regard all testing as subject to great error; and the search for objective tests largely illusory. In the final analysis the only test I really trust is provided by continuous evaluation by the instructor, reinforced by occasional oral examinations conducted by experts (that is, experts in oral examination). This, however, is, in our present circumstances, a utopian dream.

What else, then, can we in present-day circumstances reasonably do to make the effect of testing less destructive of the educational objectives we all share? We

The Multiple Choice Test with Partial Credit (MCTPC)

The idea of Cherkas and Roitberg, the inventors of MCTPC, is to scale the set of possible answers according to the level of understanding which their selection by the student indicates. Thus, for example, the problem might be to find $f \circ g(x)$, where $f(x) = x^2 - 4x + 1$ and $g(x) = 2x$; and the scoring policy for the solutions offered might be

Solution	Partial Credit	Rationale
(a) $2x^2 - 8x + 2$	2 points	correct answer for $g \circ f(x)$
(b) $4x^2 - 8x + 1$	5 points	correct answer
(c) $2x^3 - 8x^2 + 2x$	1 point	correct answer for $f(x)g(x)$
(d) $2x^2 - 8x + 1$	3 points	correct interpretation, but wrong calculation

In the experiment, each student, after taking the test, received an individualized print-out 'showing, for each question, the student's answer alongside the correct answer from {a, b, c, d}, the designation "correct", "partially correct", or "incorrect", and the corresponding point value assigned.'

must, of course, try to teach as well as possible, and we must not be content with only reaching our best students. We must establish that we and the students are on the same side and not in an adversarial situation—that we want the students to succeed. We must avoid tricks and traps; for, unfortunately, many students, basing themselves on their previous experience, are convinced that we—and the textbook writer—strew their path with hidden pitfalls, and that no question is ever as straightforward as it seems.

Of especial importance is the principle that we must not allow the tests to drive the curriculum—we must teach what we think it important for the students to know and understand, not what we can easily test. For if we teach for tests, then, inevitably, we conduct a skill-oriented course and we cannot claim that we are teaching for genuine understanding and imparting knowledge which the students will then be able to use and apply.

These recipes for reducing the distorting effect which tests—and, more generally, the students' natural emphasis on grade-acquisition—have on the whole educational enterprise contain no surprises and are only expected to achieve, at best, modest gains. But the case will have been worth making if it helps us more clearly to appreciate what we're trying to do and to realize what we're up against.

REFERENCE

- [CR] Barry Cherkas and Joseph Roitberg, Humanizing the Multiple Choice Test with Partial Credit (available from Department of Mathematics and Statistics, Hunter College of the City University of New York, 695 Park Avenue, New York, NY 10021).

*Department of Mathematical Sciences
SUNY Binghamton
Binghamton, New York 13902-6000*

A Really Trivial Proof of the Lucas-Lehmer Test

J. W. Bruce

In the paper [1] Rosen gave a beautiful and elementary proof of the Lucas-Lehmer primality test for Mersenne numbers, i.e. those of the form $M_p = 2^p - 1$. This test is very attractive, since it can be easily implemented on a computer, and it gives students the chance of experiencing the thrill of discovering very large (if well-known) primes for themselves. In this article we show that it is possible to simplify Rosen's proof of the sufficiency of the test a little to eliminate any mention of algebraic numbers or questions concerning splittings of primes. In fact we only use a couple of lemmas from group theory, which are hardly worth glorifying with such a description. Of course it is the sufficiency which allows one to produce the large primes, so the half of the result we prove has its attractions. The necessity requires some basic facts concerning quadratic residues. The proof below might make a pleasant application early on in an abstract algebra course. The author was in receipt of a Fulbright award while this note was written.

We start by defining a sequence S_n by setting $S_1 = 4$ and $S_n = S_{n-1}^2 - 2$.

Theorem 1 (LUCAS-LEHMER). *Let p be a prime number. Then $M_p = 2^p - 1$ is a prime if M_p divides S_{p-1} .*

We now follow Rosen by setting $\omega = 2 + \sqrt{3}$, $\bar{\omega} = 2 - \sqrt{3}$. One then checks that $\omega\bar{\omega} = 1$ and the next result follows easily by induction:

Lemma 1. $S_m = \omega^{2^{m-1}} + \bar{\omega}^{2^{m-1}}$.

It follows from the lemma that if M_p divides S_{p-1} then $\omega^{2^{p-2}} + \bar{\omega}^{2^{p-2}} \equiv 0 \pmod{M_p}$. Actually we will want to spell this out explicitly, so we write $\omega^{2^{p-2}} + \bar{\omega}^{2^{p-2}} = RM_p$ for some integer R . Multiplying this identity by $\omega^{2^{p-2}}$ we find that

$$\omega^{2^{p-1}} = RM_p \omega^{2^{p-2}} - 1 \quad (*)$$

and squaring

$$\omega^{2^p} = (RM_p \omega^{2^{p-2}} - 1)^2 \quad (**).$$

Now assume that M_p is composite, and choose a prime divisor q with $q^2 \leq M_p$. (This is where our proof departs from that of Rosen.) At one stage below we need the fact that q obviously is not 2.

We are going to use some really elementary results and ideas from group theory.

Lemma 2. *Let X be a set with a binary operation which is associative and has an identity. Then the set X^* of invertible elements in X forms a group.*

Proof: Clearly the identity $1 \in X^*$, so we have a non-empty set. We now have only to show that the set X^* is closed under the binary operation. But if x_1 and x_2 are invertible elements with inverses x_1^{-1} , x_2^{-1} then $x_1 x_2$ has inverse $x_2^{-1} x_1^{-1}$.

Lemma 3. *If G is a finite group then the order of an element is at most the order of the group. If $x \in G$ and $x^r = 1$ then the order of x divides r .*

(Results in group theory do not get any easier than this. These results are nowhere near as deep as Lagrange's Theorem for example!)

Proof of Theorem 1: Let Z_q denote the set of integers modulo q , and X denote the set $\{a + b\sqrt{3} : a, b \in Z_q\}$. We can define two binary operations on X , namely addition and multiplication, in the obvious manner. So in the case of multiplication, which is the one of interest, we choose representatives in $Z[\sqrt{3}]$ of our elements of X compute the product in the usual way, $(a_1 + b_1\sqrt{3})(a_2 + b_2\sqrt{3})$ as $(a_1 a_2 + 3b_1 b_2) + (a_1 b_2 + a_2 b_1)\sqrt{3}$ and then reduce the coefficients modulo q . In the case of addition we obviously get an abelian group, and for multiplication we clearly have an associative (and commutative) binary operation with identity 1. Let X^* denote the group of invertible elements of X with respect to multiplication. Lemma 2 tells us that this is a group, while Lemma 3 tells us that the order of any element of X^* is at most $q^2 - 1$, since X^* contains at least one non-invertible element, namely 0.

Now consider $\omega = 2 + \sqrt{3}$ as an element of X . Since q divides M_p it follows that $RM_p \omega^{2^{p-2}}$, when viewed as an element of X , is 0. So the equalities noted in (*) and (**) above in X reduce to $\omega^{2^{p-1}} = -1$, and $\omega^{2^p} = 1$ respectively. It follows that ω lies in X^* , and has order 2^p . For the order of ω clearly divides 2^p by Lemma 3 and the second equality, but cannot be less than 2^p by the first. So using Lemma 3 again we deduce that $2^p \leq q^2 - 1$. However $q^2 - 1 \leq M_p - 1 = 2^p - 2$ and we have a contradiction.

REFERENCE

1. M. I. Rosen, A Proof of the Lucas-Lehmer Test, Amer. Math. Monthly, 95 (1988), 855–856.

*Department of Pure Mathematics
The University of Liverpool
L69 3BX, UNITED KINGDOM*

Pascal's Matrices

Gregory S. Call and Daniel J. Velleman

We first encountered Pascal's matrices while working on a probability problem involving repeated flips of an unfair coin. Since they are derived naturally from Pascal's triangle, Pascal's matrices immediately caught our attention. However, it was the striking simplicity of the powers of these matrices that intrigued us and prompted us to write this paper. Before defining Pascal's matrices, we'll describe the probability question which led us to them.

If the probability of heads on each flip is p , and thus the probability of tails is $1 - p$, then in a sequence of n flips the probability of any event is given by an expression of the form

$$\sum_{i=0}^n a_i p^{n-i} (1-p)^i,$$

where the a_i 's are integers which depend on the event in question, with $0 \leq a_i \leq \binom{n}{i}$. Of course, this can be multiplied out to yield a polynomial in p of degree n with integer coefficients. While trying to resolve an open problem in [1], we found that we needed to reverse this process: Given a polynomial in p with integer coefficients, is it the formula for the probability of some event? As we will show later, the answer to this question involves a Pascal's matrix.

The $n \times n$ Pascal's matrix is obtained by taking the first n rows of Pascal's triangle and filling in with 0's on the right. Specifically, we define the $n \times n$ Pascal's matrix P by

$$P_{ij} = \begin{cases} \binom{i-1}{j-1} & \text{if } i \geq j, \\ 0 & \text{otherwise.} \end{cases}$$

Thus the first four Pascal's matrices are

$$(1), \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{pmatrix}.$$

Computing the inverses of these matrices reveals a rather suggestive pattern. In particular, for $n = 4$ we have

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ -1 & 3 & -3 & 1 \end{pmatrix}.$$

These examples lead one to conjecture that the inverse of a Pascal's matrix is obtained by multiplying every other (nonzero) entry by -1 . As we show below, this conjecture may be proved using the binomial theorem.

Theorem 1. Let Q be the $n \times n$ matrix defined by

$$Q_{ij} = \begin{cases} (-1)^{i-j} \binom{i-1}{j-1} & \text{if } i \geq j, \\ 0 & \text{otherwise.} \end{cases}$$

Then $P^{-1} = Q$.

Proof: It is clear that $(PQ)_{ij} = \begin{cases} 0 & \text{if } i < j, \\ 1 & \text{if } i = j. \end{cases}$ Hence, it suffices to show that $(PQ)_{ij} = 0$ if $i > j$. Suppose $i > j$ and write $i = j + l$ with $l > 0$. Then

$$\begin{aligned} (PQ)_{ij} &= \sum_{k=0}^l P_{j+l, j+k} Q_{j+k, j} = \sum_{k=0}^l \binom{j+l-1}{j+k-1} \binom{j+k-1}{j-1} (-1)^k \\ &= \sum_{k=0}^l \frac{(j+l-1)!}{(l-k)!(j-1)!k!} (-1)^k = \frac{(j+l-1)!}{(j-1)!l!} \sum_{k=0}^l \frac{l!}{(l-k)!k!} (-1)^k \\ &= \binom{j+l-1}{j-1} \sum_{k=0}^l \binom{l}{k} (-1)^k = \binom{j+l-1}{j-1} (1-1)^l = 0. \end{aligned}$$

The key idea in the preceding proof is to use the binomial theorem to isolate the expansion of $(1-1)^l$. Since the binomial theorem gives a formula for the expansion of $(x+y)^l$ for any real numbers x and y , Theorem 1 has a natural generalization. For any nonzero real number x , define

$$P[x] = \begin{cases} x^{i-j} \binom{i-1}{j-1} & \text{if } i \geq j, \\ 0 & \text{otherwise,} \end{cases}$$

and let $P[0]$ equal the identity matrix. Observe that if we were to adopt the convention that $0^0 = 1$, then the above formula for $P[x]$ would also yield $P[0]$ equals the identity. Furthermore, notice that $P[1] = P$ and, by Theorem 1, $P[-1] = P^{-1}$. For example, for $n = 4$ we have

$$P[x] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ x & 1 & 0 & 0 \\ x^2 & 2x & 1 & 0 \\ x^3 & 3x^2 & 3x & 1 \end{pmatrix}.$$

Theorem 2. For any real numbers x and y , $P[x]P[y] = P[x+y]$.

Proof: If $x = 0$ or $y = 0$, the assertion is trivial. If neither x nor y is zero, then the proof is similar to Theorem 1. In particular, one may use the binomial theorem to isolate the expansion of $(x+y)^l$.

Corollary 3. For any integers j and k , $k \neq 0$,

- (a) $P^j = P[j]$,
- (b) $(P[j/k])^k = P^j$.

Proof: Recall $P[1] = P$ and $P[0]$ is the identity matrix. Hence, it follows from Theorem 2, by induction on j , that $P[j] = P^j$ for all $j \geq 0$. Since $P[-1] = P^{-1}$, a

similar induction on $|j|$ shows $P[j] = P^j$ for all $j < 0$. Then, by Theorem 2 and (a), we have $(P[j/k])^k = P[j] = P^j$.

Corollary 3(b) shows that for any rational number x , we may regard $P[x]$ as the “ x^{th} power of P .” For example, for $n = 4$, we have square and cube roots of P given by:

$$P\left[\frac{1}{2}\right] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{4} & 1 & 1 & 0 \\ \frac{1}{8} & \frac{3}{4} & \frac{3}{2} & 1 \end{pmatrix}, \quad P\left[\frac{1}{3}\right] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{3} & 1 & 0 & 0 \\ \frac{1}{9} & \frac{2}{3} & 1 & 0 \\ \frac{1}{27} & \frac{1}{3} & 1 & 1 \end{pmatrix}.$$

If x is irrational, then does $P[x]$ still represent a matrix which deserves to be regarded as “ P^x ”? Recall how irrational exponents for real numbers work. If $a > 0$, then a^x is defined to be e^{xl} , where $l = \ln a$. By analogy, if $P[x]$ is to be regarded as “ P^x ”, we might expect that $P[x] = e^{xL}$, for some matrix L . Is there such a matrix?

Notice that we are applying the exponential function to a matrix here. Matrix exponentials are defined by simply plugging matrices into the usual Maclaurin series for the exponential function. In other words, for any square matrix A , the exponential of A is defined to be the matrix

$$e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots + \frac{A^k}{k!} + \cdots$$

It can be shown that this series converges for every square matrix A , in the sense that each entry in the matrix series converges. Furthermore, the matrix exponential has many of the same properties as the ordinary exponential function, as the following theorem shows. For proofs of these facts, see [2, p. 518].

Theorem 4. *Let A be any square matrix. Then:*

- (a) *For any numbers s and t , $e^{(s+t)A} = e^{sA}e^{tA}$.*
- (b) *e^A is invertible, and $(e^A)^{-1} = e^{-A}$.*
- (c) *$\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A$, where $\frac{d}{dt}e^{tA}$ is the matrix resulting from taking the derivative with respect to t of each entry of e^{tA} .*

Suppose there is a matrix L such that $P[x] = e^{xL}$. Then $\frac{d}{dx}P[x] = Le^{xL} = LP[x]$, so $\frac{d}{dx}P[x]|_{x=0} = LP[0] = LI = L$. Thus, there is at most one matrix L such that $P[x] = e^{xL}$. For example, in the case $n = 4$ we can find the only possible value for L as follows:

$$\frac{d}{dx}P[x] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2x & 2 & 0 & 0 \\ 3x^2 & 6x & 3 & 0 \end{pmatrix}, \quad L = \frac{d}{dx}P[x]|_{x=0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

This suggests how we should choose L in general. Let L be the $n \times n$ matrix with entries

$$L_{ij} = \begin{cases} j & \text{if } i = j + 1, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 5. For every real number x , $P[x] = e^{xL}$.

Note that in particular, taking $x = 1$ in Theorem 5, we have $P = e^L$. To prove Theorem 5, we will need the following lemma.

Lemma 6. For every positive integer k , the entries of L^k are given by the formula

$$(L^k)_{ij} = \begin{cases} \frac{(i-1)!}{(j-1)!} & \text{if } i = j + k, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: By induction on k . Note that for $k \geq n$ we have $L^k = 0$.

Proof of Theorem 5: Clearly if $x = 0$ then $e^{xL} = I = P[x]$. Now suppose $x \neq 0$. Then since $L^k = 0$ for $k \geq n$, the infinite series for e^{xL} reduces to the finite sum

$$e^{xL} = I + xL + \frac{x^2}{2}L^2 + \cdots + \frac{x^{n-1}}{(n-1)!}L^{n-1}.$$

Applying Lemma 6, we can now read off the entries in e^{xL} . Clearly it is a lower triangular matrix, and the diagonal entries are all 1. Now suppose $i > j$, and let $k = i - j$. Then the only matrix in the sum above which has a nonzero ij -entry is $(x^k/k!)L^k$, so

$$(e^{xL})_{ij} = \frac{x^k}{k!}(L^k)_{ij} = x^k \frac{(i-1)!}{(j-1)!(i-j)!} = x^k \binom{i-1}{j-1} = (P[x])_{ij}.$$

Using Theorems 4 and 5 we can now give a one line proof of Theorem 2:

$$P[x]P[y] = e^{xL}e^{yL} = e^{(x+y)L} = P[x+y].$$

Now let's return to the problem involving repeated flips of an unfair coin described in the second paragraph. If we start with the formula for the probability of an event, and expand $(1-p)^i$ by the binomial theorem, we find that

$$\begin{aligned} \sum_{i=0}^n a_i p^{n-i} (1-p)^i &= \sum_{i=0}^n a_i p^{n-i} \sum_{j=0}^i \binom{i}{j} (-p)^{i-j} \\ &= \sum_{i=0}^n \sum_{j=0}^i p^{n-j} a_i \binom{i}{j} (-1)^{i-j} = \sum_{j=0}^n \sum_{i=j}^n p^{n-j} a_i \binom{i}{j} (-1)^{i-j}. \end{aligned}$$

Thus, if we let

$$b_j = \sum_{i=j}^n a_i \binom{i}{j} (-1)^{i-j}$$

then we have

$$\sum_{i=0}^n a_i p^{n-i} (1-p)^i = \sum_{j=0}^n b_j p^{n-j}.$$

The relationship between the a_i 's and the b_j 's turns out to involve the $(n+1) \times (n+1)$ Pascal's matrix! To see this, let \vec{a} and \vec{b} be the row vectors

$$\vec{a} = (a_0, a_1, \dots, a_n), \quad \vec{b} = (b_0, b_1, \dots, b_n).$$

Then the definition of the b_j 's simply says

$$\vec{b} = \vec{a}P[-1] = \vec{a}P^{-1}.$$

Thus, given a vector \vec{b} of integers which are the coefficients of a polynomial in p of degree n , we can recover the vector of coefficients \vec{a} in the formula for the probability of an event by simply multiplying by Pascal's matrix:

$$\vec{a} = \vec{b}P.$$

Once we have computed \vec{a} , we can tell if the original polynomial was the formula for the probability of some event by simply checking whether or not $0 \leq a_i \leq \binom{n}{i}$ for $0 \leq i \leq n$.

REFERENCES

1. I. Szalkai and D. Velleman, Versatile coins, *American Mathematical Monthly* 100 (1993) 26–33.
2. R. Williamson and H. Trotter, *Multivariable Mathematics*, second edition, Prentice-Hall, 1979.

Department of Mathematics and Computer Science
Amherst College
Amherst, MA 01002
gscall@amherst.edu
djvelleman@amherst.edu

Throughout the 1960s and 1970s devoted Beckett readers greeted each successively shorter volume from the master with a mixture of awe and apprehensiveness; it was like watching a great mathematician wielding an infinitesimal calculus, his equations approaching nearer and still nearer to the null point.

The writer is *John Banville*. The essay from which this gem is excerpted is entitled “The Last Word”, which is a review of Samuel Beckett’s *Nohow On: Company, III Seen III Said, Worstward Ho*. It appeared in the August 13, 1992 issue of *The New York Review of Books*.

—submitted by Bill Rosenthal
 College of Education
 Michigan State University
 East Lansing, MI 48824

Symmetries of the Cube and Outer Automorphisms of S_6

Thomas A. Fournelle

Let S_n denote the group of all permutations of the set $\{1, 2, \dots, n\}$. Each element a of a group G induces an automorphism $(g)\delta_a = a^{-1}ga$. Such an automorphism is called *inner*, otherwise an automorphism is called *outer*. It is well known [4, Theorem 7.4] that S_n has no outer automorphisms if $n \neq 6$. Hölder [1] proved in 1895 that S_6 has an outer automorphism. The article of Janusz and Rotman [2] is an extremely well written account of various proofs given since 1895 of the existence of an outer automorphism of S_6 . That article also gives an explicit construction of an outer automorphism of S_6 of order 2 and a proof that the automorphism group of S_6 is generated by this outer automorphism together with all inner automorphisms. (A simplification of the Janusz-Rotman article appears in [5].)

The purpose of this article is to present a heuristic argument that suggests the possibility of an outer automorphism of S_6 . This argument will be pursued until it yields an explicit construction of an outer automorphism. The construction here will be quite a bit different from those given previously.

To begin, recall that an *isometry* of \mathbb{R}^3 is a bijection which preserves distance. The set of all isometries of \mathbb{R}^3 forms a group under composition of functions. An isometry which fixes the origin can be identified with a 3×3 orthogonal matrix. Recall that a matrix X is orthogonal if its transpose is its inverse. Such a matrix, of course, has determinant ± 1 . The set of all such matrices forms the *orthogonal group* $O(3, \mathbb{R})$. Now consider a cube in \mathbb{R}^3 centered at the origin and positioned so that the x , y , and z -axes, with right hand orientation, go through the centers of the 6 faces. Label the positive and negative axes with the numbers 1, 2, \dots , 6 as in figure 1. Let H be the group of isometries on the cube. Our goal is to show that

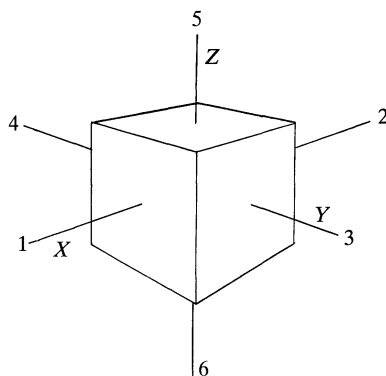


Figure 1

the group S_6 has two distinct subgroups isomorphic to H and that S_6 has an outer automorphism which interchanges these two subgroups. The first of these subgroups is easy to see by considering figure 1. Obviously, an element of H is determined by its action on $1, 2, \dots, 6$. Hence, H may be considered a subgroup of S_6 as well as $O(3, \mathbb{R})$.

The isometries of the cube are of two types, those which preserve orientation and those which do not. Those which preserve orientation have determinant 1 when considered as matrices in $O(3, \mathbb{R})$. Since the product of matrices of determinant 1 is of determinant 1, it follows that the orientation preserving elements of H form a subgroup, which will be denoted by H_0 . The cube has 4 diagonals which are permuted by the elements of H_0 . Hence, there is a homomorphism

$$\sigma: H_0 \rightarrow S_4.$$

If the diagonals of the cube are labeled as in figure 2 then the values of σ can be computed explicitly. For example, the rotation around the x -axis represented by the permutation (3546) in H_0 induces the permutation (1243) on the diagonals. Thus, it follows that

$$(3546)^\sigma = (1243). \tag{1}$$

Similarly, by considering a rotation around diagonal 1, one obtains

$$((145)(236))^\sigma = (243). \tag{2}$$

Now $(1243)(243)(243) = (13)$. It is a standard (easy) exercise show that (13) and (1243) generate all of S_4 . Hence, σ is surjective. Hence, $|H_0| \geq o(S_4) = 24$. A simple counting argument shows that $|H| = 3!2^3 = 48$. Since H_0 is a proper subgroup of H it follows that $|H_0| = 24$. Hence, σ is an isomorphism.

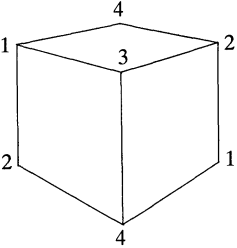


Figure 2

In H the element $i = (12)(34)(56)$ represents the inversion through the origin. Clearly, i has order 2. Perhaps less clear is the fact that i is central, that is, commutes with every element of H . This may be seen most easily by considering the isometries of the cube to be elements of the orthogonal group on \mathbb{R}^3 . In this matrix group i is represented by

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

which is clearly central. Now H_0 is normal in H since it is of index 2 and $\langle i \rangle$ is normal since it is central. Since H_0 and $\langle i \rangle$ intersect trivially it follows that

$$H = H_0 \times \langle i \rangle \cong S_4 \times C_2.$$

Now consider the graphs in figures 3 and 4. Now the isometries in H are determined by the images of $1, 2, \dots, 6$ in figure 1, and these symmetries must send the various axes to axes. It follows by examining figure 3 that the group of graph automorphisms of the graph is isomorphic to H . (Incidentally, the group of symmetries of an n -dimensional cube is isomorphic to the group of graph automorphisms of a graph similar to figure 3. The graph corresponding to the 4-dimensional cube is given in figure 5. It is clearly easier to draw this planar graph than to draw the 4-dimensional cube!) Now let K be the group of graph automorphisms of figure 4. Clearly, $K \cong S_4 \times C_2$, and K may be considered a subgroup of S_6 . We can consider the S_4 factor of K to be the image of H_0 under σ . Since the C_2 factor of H is generated by $i = (12)(34)(56)$ we may extend σ to an isomorphism

$$\sigma: H \rightarrow K$$

by defining

$$((12)(34)(56))^\sigma = (56). \quad (3)$$

H and K are both subgroups of S_6 and it is natural to ask whether σ extends to an automorphism of S_6 . Suppose it does. It is fundamental that inner automorphisms of S_n preserve cycle structure. From (3) it follows that if σ does extend to an automorphism of S_6 then it must be an outer automorphism.

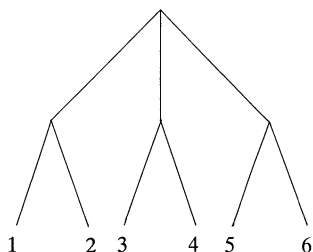


Figure 3

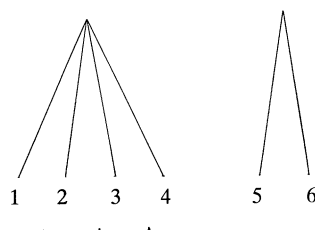


Figure 4

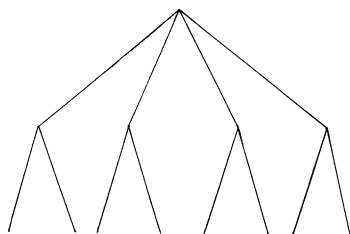


Figure 5

It is now imperative to find a proof that σ extends to an automorphism of S_6 . There are (at least) four ways to do this:

1. The author used the group theoretic programming language CAYLEY to construct the automorphism group of S_6 and then check to see if any automorphisms match the images of σ given in (1), (2), and (3). Since (3546) , $(145)(236)$ and $(12)(34)(56)$ generate H , any automorphism matching (1), (2), and (3) must be an extension of σ . If this is done it will be found that σ extends to a number of automorphisms, exactly one of which has order 2.

2. In [3] Miller constructs an outer automorphism of S_6 he calls φ which is defined by

$$(12)^\psi = (12)(36)(45),$$

$$(13)^\psi = (13)(24)(56),$$

$$(14)^\psi = (14)(26)(35),$$

$$(15)^\psi = (15)(23)(46),$$

$$(16)^\psi = (16)(25)(34).$$

Now let τ be the inner automorphism of S_6 induced by (1356)(24). Then routine calculation shows that

$$(3546)^{\tau\psi\tau^{-1}} = (1243), \quad (4)$$

$$((145)(236))^{\tau\psi\tau^{-1}} = (243), \quad (5)$$

$$((12)(34)(56))^{\tau\psi\tau^{-1}} = (56). \quad (6)$$

Thus, σ extends to the automorphism $\tau\psi\tau^{-1}$. Incidentally, ψ (and hence $\tau\psi\tau^{-1}$) is of order 2.

3. In [2] Janusz and Rotman construct an outer automorphism φ of S_6 of order 10. With their Lemma 6 and Corollary 8 it can be shown as above that $\tau\varphi^5\tau^{-1}$ is an extension of σ for appropriately chosen τ . (Note that in this paper functions act on the right. Hence, $(12)(13) = (123)$. In [2] functions act on the left so that $(12)(13) = (132)$. The notation in this paper is consistent with CAYLEY and with [3].)

4. If one assumes that σ extends to an automorphism of order 2 then values of σ can be computed. For example, since σ is of order 2 equation (3) becomes

$$(56)^\sigma = (12)(34)(56).$$

One can compute the values of σ on a set of generators of S_6 assuming that σ extends to an automorphism of order 2. Then by checking that σ preserves the relations among the generators it can be seen that σ actually does extend to an automorphism of order 2. This method of showing that σ extends to an automorphism of S_6 is a bit more tedious than the other methods mentioned here but it can be done by hand and does not need to refer to the constructions in (2) or (3).

REFERENCES

1. O. Holder, Bildung zusammengesetzter Gruppen, *Math. Ann.* 46 (1895), 321–422.
2. G. Janusz and J. J. Rotman, Outer automorphisms of S_6 , this MONTHLY, 89 (1982), 407–410.
3. D. W. Miller, On a theorem of Hölder, this MONTHLY, 65 (1958), 252–254.
4. J. J. Rotman, *The Theory of Groups, An Introduction*, 2nd ed., Allyn and Bacon, Boston, MA, 1973.
5. Advanced Problem 6538, this MONTHLY, 95 (1988), 779.

*Department of Mathematics
University of Wisconsin-Parkside
Kenosha, WI 53141*

Isogonal Configurations

Timothy A. Murdoch

A familiar and useful fact from linear algebra is that every non-zero subspace of an inner product space \mathbf{V} is spanned by an orthogonal set of vectors. Moreover, the familiar Gram-Schmidt orthogonalization procedure shows how to construct such orthogonal spanning sets (see [HK], p. 280). However, when considering the geometry of \mathbf{V} , a natural question arises about the existence of other geometrically interesting configurations of vectors that are analogous to orthogonal configurations. In this article, we show that for finite dimensional inner product spaces, orthogonal sets of vectors may be considered as a special case of *isogonal*, or equiangular, sets of vectors. That is, we show that for angles φ strictly between 0 and $\arccos(-1/(n-1))$, where n is the dimension of \mathbf{V} , there exist sets of n linearly independent vectors in \mathbf{V} with the property that the angle between any two elements in the set is φ . In fact, only for angles strictly between 0 and $\arccos(-1/(n-1))$ do there exist isogonal configurations of n linearly independent vectors. For these allowable angles, we give an analogue of the Gram-Schmidt orthogonalization procedure for constructing a φ -isogonal spanning set of vectors from a given orthonormal basis. Perhaps somewhat surprising is our proof that there is a unique isogonal configuration (up to rigid motion) of non-zero linearly dependent vectors whose span is \mathbf{V} . We wish to thank the referee for suggestions improving the organization of the results.

Before proceeding to the discussion of the general case, the reader may want to consider first the cases when $n = 2$ and $n = 3$. When $n = 2$, there is a basis for \mathbf{R}^2 such that the angle between the basis vectors is any given angle between zero and π . The linearly dependent configuration satisfying the angle condition must be the vertices of an equilateral triangle. When $n = 3$, the bases of \mathbf{R}^3 such that the angle made by each pair of vectors in the basis is the same is only possible for angles between zero and $2\pi/3$. The linearly dependent case gives the vertices of a regular tetrahedron.

Our interest in this construction was stimulated by the appearance of the linearly dependent configurations in \mathbf{R}^2 and \mathbf{R}^3 in the structure of soap films and soap bubbles in ordinary 3-space. These cases are pictured in the articles by Almgren and Taylor [AT] and Hildebrandt [H].

To have a concrete model for the general construction, we let $\mathbf{V} = \mathbf{R}^n$ equipped with the standard dot product. Our goal is to understand the configurations in \mathbf{R}^n which generalize the familiar cases in \mathbf{R}^2 and \mathbf{R}^3 .

Definition. A set $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ of unit vectors in \mathbf{R}^n is said to be φ -isogonal if $\mathbf{x}_i \cdot \mathbf{x}_j = \cos \varphi$ for all i, j with $i \neq j$.

Of course, for vectors of the same length the condition that the set be φ -isogonal is equivalent to the condition that the distance between any pair of vectors, considered as points on the unit sphere, is the same.

Theorem 1. Let $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ be a φ -isogonal set with $\varphi \neq 0$. Let d be the dimension of the subspace spanned by S . Then either:

(i) $d = m$ and $0 < \varphi < \arccos(-1/(d-1))$,

or

(ii) $d = m - 1$ and $\varphi = \arccos(-1/d)$. In this case, $\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_m = \mathbf{0}$ and the distance between any two points of the configuration is $\sqrt{2 + 2/d}$.

Proof: Let \mathbf{G} be the Gram matrix formed from the vectors in S ; i.e., the symmetric $m \times m$ matrix with i, j -entry $\mathbf{x}_i \cdot \mathbf{x}_j$. Recall the following facts about the Gram matrix \mathbf{G} :

(a) \mathbf{G} is positive semi-definite,

(b) \mathbf{G} is positive definite if and only if S is linearly independent,

(c) the eigenvalues of \mathbf{G} are non-negative, and

(d) the eigenvalues of \mathbf{G} are strictly positive if and only if S is linearly independent.

We determine the eigenvalues of \mathbf{G} by letting \mathbf{I} denote the $m \times m$ identity matrix and letting \mathbf{J} denote the $m \times m$ matrix with a 1 in each entry. Then $\mathbf{G} = (1 - \cos \varphi)\mathbf{I} + \cos \varphi \mathbf{J}$, so the eigenvalues of \mathbf{G} are $\lambda_1 = 1 + (m-1)\cos \varphi$, of multiplicity one, and $\lambda_2 = 1 - \cos \varphi$, of multiplicity $m-1$. Since $\varphi \neq 0$, we have $\lambda_2 > 0$. Since \mathbf{G} is a positive semi-definite matrix, we must have $\lambda_1 \geq 0$, so $\cos \varphi \geq -1/(m-1)$.

If $d = m$, then S is a linearly independent set and \mathbf{G} is a positive definite matrix. Thus $\lambda_1 > 0$ and $\cos \varphi > -1/(m-1)$, establishing case (i).

If $d < m$, then S is a linearly dependent set, so \mathbf{G} is a singular matrix and $\lambda_1 = 0$. Hence $\varphi = \arccos(-1/(m-1))$. Since d is the dimension of the subspace spanned by S , this argument applies to any subset of $d+1$ vectors of S . Therefore, we must also have $\varphi = \arccos(-1/d)$, and thus $d = m-1$.

Now, the eigenvalue λ_1 has associated eigenvector $\mathbf{e} = (1, 1, \dots, 1)^t$, so when $d < m$ we have $\mathbf{G}\mathbf{e} = \mathbf{0}$. Hence $\mathbf{e}^t \mathbf{G} \mathbf{e} = 0$. However,

$$\mathbf{e}^t \mathbf{G} \mathbf{e} = \sum_{j=1}^m \sum_{i=1}^m \mathbf{x}_i \cdot \mathbf{x}_j = \|\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_m\|^2,$$

so $\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_m = \mathbf{0}$.

Finally, compute

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = (\mathbf{x}_i - \mathbf{x}_j) \cdot (\mathbf{x}_i - \mathbf{x}_j) = 2 - 2\cos \varphi$$

to show that the distance between any two points in S is $\sqrt{2 + 2/d}$ to establish case (ii) and complete the proof of the theorem.

Corollary. If $s = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is a φ -isogonal set which spans \mathbf{R}^n then either

(i) $m = n$ and $0 < \varphi < \arccos(-1/(n-1))$,

or

(ii) $m = n + 1$ and $\varphi = \arccos(-1/n)$.

Remarks. (a) The expression for the largest angle permitting an isogonal configuration in \mathbf{R}^n , $\varphi_{\max} = \arccos(-1/n)$, decreases as n increases.

(b) The corollary shows that there is exactly one linearly dependent isogonal configuration of vectors (up to rigid motion) which spans \mathbf{R}^n —the vertices of a regular n -simplex. Thus, no isogonal configuration exists for $n+2$ or more vectors in \mathbf{R}^n .

Now suppose S is a basis for \mathbf{R}^n . Using the Gram-Schmidt orthogonalization procedure (if necessary), we assume further that S consists of an orthonormal basis $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$. We construct, by induction, an isogonal basis (of unit length vectors) $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ for \mathbf{R}^n such that the angle φ between any two vectors is a given angle (strictly) between zero and $\arccos(-1/(n-1))$.

Let $\mathbf{z}_1 = \mathbf{y}_1$, and define $\mathbf{z}_2 = \cos \varphi \mathbf{z}_1 + \sin \varphi \mathbf{y}_2$. Then $\{\mathbf{z}_1, \mathbf{z}_2\}$ is easily checked to be a φ -isogonal set such that $\text{span}\{\mathbf{z}_1, \mathbf{z}_2\} = \text{span}\{\mathbf{y}_1, \mathbf{y}_2\}$. In general, given the set of φ -isogonal vectors $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{j-1}$, define

$$\mathbf{z}_j = \frac{\cos \varphi}{1 + (j-2)\cos \varphi} (\mathbf{z}_1 + \mathbf{z}_2 + \dots + \mathbf{z}_{j-1}) + \gamma_j \mathbf{y}_j \quad (*)$$

where $\gamma_j = \sqrt{((1 - \cos \varphi)(1 + (j-1)\cos \varphi)) / (1 + (j-2)\cos \varphi)}$. Then $\mathbf{z}_j \in \text{span}\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_j\}$, so that $\mathbf{z}_j \cdot \mathbf{y}_i = 0$ for $i \geq j+1$. Moreover, by using (*) it is straightforward to check that $\text{span}\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_j\} = \text{span}\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_j\}$ for $j = 1, 2, \dots, n$, and that the \mathbf{z}_j 's constitute a φ -isogonal set of unit vectors. Also, note that if $\varphi = \pi/2$, then $\mathbf{z}_j = \mathbf{y}_j$ for each $j = 1, 2, \dots, n$; i.e., unit length $\pi/2$ -isogonal vectors are, of course, just orthonormal vectors.

Remarks. (a) The formula (*) for the \mathbf{z}_j shows that it is possible to construct φ -isogonal sets of vectors in an inner product space with a countable basis provided that $0 < \varphi \leq \pi/2$, the case $\varphi = \pi/2$ being, of course, orthogonal sets.

(b) The formula (*) for the \mathbf{z}_j also shows that all φ -isogonal sets of unit vectors are equivalent up to rigid motion. This follows from the construction of the isogonal vectors as linear combinations of the orthonormal vectors and the inner product preserving property of rigid motions.

Examining the geometry underlying the construction of φ -isogonal configurations makes the relationship between the angle values and the existence (or non-existence) of the configurations more transparent. The φ -isogonal vectors are constructed by first taking the cone of vectors C_1 which make angle φ with the vector \mathbf{y}_1 of the given orthonormal set and choosing a vector \mathbf{z}_2 in this cone. Next, take the cone of vectors C_2 which make angle φ with \mathbf{z}_2 and choose a vector \mathbf{z}_3 in the intersection of C_1 and C_2 (see Figure 1). Continuing this process, the cones about the vectors $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_j$ will intersect as long as the angle φ lies between zero and $\arccos(-1/(j-1))$. When $\varphi = \arccos(-1/(j-1))$, each set of $j-1$ cones intersect along a unique line segment and case (ii) of the theorem states that

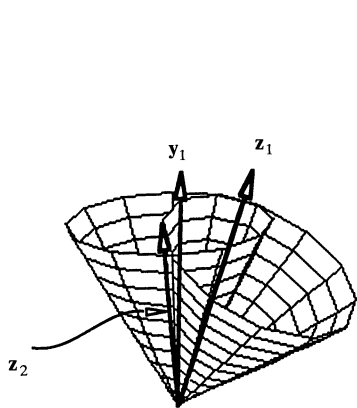


Figure 1

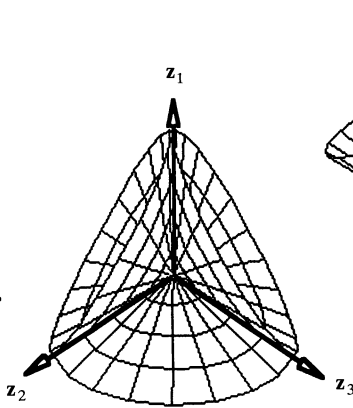


Figure 2

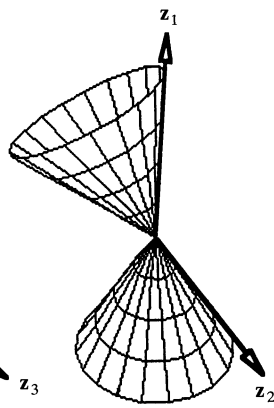


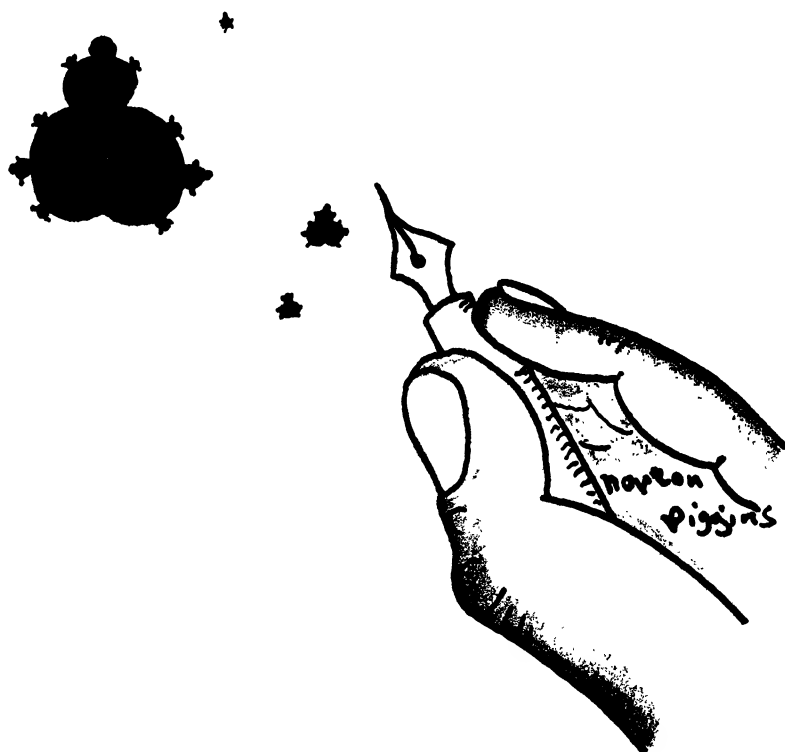
Figure 3

the resulting configuration lies in a proper subspace of dimension $j - 1$ (see Figure 2 for the case when $n = j = 3$). Moreover, if $\cos \varphi < -1/(j - 1)$, then the cone about \mathbf{z}_{j-2} , C_{j-2} , can intersect the cones C_1, C_1, \dots, C_{j-1} at most in the vectors $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{j-2}$ (see Figure 3 for the case $n = 2, j = 3$).

REFERENCES

- [AT] Almgren, F. and Taylor, J., *The Geometry of Soap Bubbles and Soap Films*, Scientific American, July 1976.
 [H] Hildebrandt, S., *The Calculus of Variations Today*, Mathematical Intelligencer, Fall 1989.
 [HK] Hoffman, K. and Kunze, R., *Linear Algebra*, 2nd ed., Prentice Hall 1971.

Mathematics Department
Washington and Lee University
Lexington, VA 24450



A MANDEL BLOT

NOTES

Edited by: John Duncan

$$PSL_2(\mathbf{Z}) = \mathbf{Z}_2 * \mathbf{Z}_3$$

Roger C. Alperin

We shall prove that the modular group $\Gamma = PSL_2(\mathbf{Z})$ has the structure of a free product of a cyclic group of order 2 and a cyclic group of order 3. Usually this result is obtained by finding a fundamental domain for the action on the upper half-plane. Here the result is proved in a quite surprising manner using only the action on the irrational numbers. The proof uses the characterization of a free product as the set of alternating words (cf. [L-S, Proposition 12.2]): viz., G is the free product of its subgroups A and B , denoted, $G = A * B$, if and only if it is generated by these subgroups and if $w = a_1 b_1 a_2 b_2 \cdots a_n b_n$ for $a_i \in A$, $b_j \in B$, $1 \leq i, j \leq n$ and a_i, b_j are different from the identity except possibly for $i = 1$ or $j = n$, then w is not the identity element of G .

The group Γ is the quotient group $SL_2(\mathbf{Z})/\{\pm I\}$ where $SL_2(\mathbf{Z})$ is the group of 2×2 integer matrices of determinant 1. It is easy to see using row and column operations and the Euclidean algorithm that the matrices

$$\mu = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

generate $SL_2(\mathbf{Z})$ and hence their images generate Γ . Now let $\beta = \mu\alpha$ and denote similarly their images in Γ . It is clear now that Γ is generated by $A = \langle \alpha \rangle$ and $B = \langle \beta \rangle$. The subgroup A is cyclic of order 2 and the subgroup B is cyclic of order 3.

The group Γ acts via linear fractional transformations on the extended complex numbers and hence also on the set of \mathcal{S} of real irrational numbers. Explicitly if $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in $SL_2(\mathbf{Z})$ then the action on the irrationals is given by

$$z \rightarrow \frac{az + b}{cz + d}.$$

The action of the generators is given by

$$\alpha: z \rightarrow \frac{-1}{z}$$

$$\beta: z \rightarrow -\frac{1}{z}$$

and

$$\beta^{-1}: z \rightarrow \frac{1}{1-z}.$$

To obtain the theorem of the title we prove the alternating word characterization of free products; for this we make some observations about the action. Let \mathcal{P} denote the set of positive irrationals and \mathcal{N} denote the set of negative irrationals.

It is clear that

$$\alpha(\mathcal{P}) \subset \mathcal{N}$$

and

$$\beta^\pm(\mathcal{N}) \subset \mathcal{P}.$$

We are now ready to verify the alternating word property. Given a word w which is alternating from A to B , if it is of odd length as a word in α, β^\pm then either $w(\mathcal{P}) \subset \mathcal{N}$ or $w(\mathcal{N}) \subset \mathcal{P}$ depending on whether the rightmost letter is α or not. If the word is of even length, we can conjugate by α if necessary to obtain a new word w which begins with a power of β and ends with an α . Now, if $w = \beta \cdots \alpha$ then $w(\mathcal{P}) \subset \beta(\mathcal{N})$ is a subset of irrationals greater than 1; similarly, if $w = \beta^{-1} \cdots \alpha$ then $w(\mathcal{P}) \subset \beta^{-1}(\mathcal{N})$ is a subset of positive irrationals less than 1. In any case then $w(z) \neq z$ for some irrational z and thus it is not the identity; since this is a conjugate of the given word we have verified that it too is not the identity element.

REFERENCES

[L-S] Lyndon, Roger C. and Schupp, Paul E., Combinatorial Group Theory, Springer-Verlag, New York, 1971.

*Department of Mathematics and Computer Science
San Jose State University
San Jose, CA 95192*

Generators for the Algebra of Symmetric Polynomials

D. G. Mead

With n a fixed positive integer, let Σ denote the ring of symmetric polynomials in the variables x_1, x_2, \dots, x_n with rational coefficients. As is well known, this ring is generated by the elementary symmetric functions in x_1, x_2, \dots, x_n and also by the first n power symmetric functions $p_i = x_1^i + x_2^i + \cdots + x_n^i$, $1 \leq i \leq n$. In response to a question raised by S. K. Stein in a conversation, we show that these facts are two cases of a more general theorem concerning families that generate Σ .

First, a few definitions. Let $Q[x_1, x_2, \dots, x_n]$ be the ring of polynomials in the variables x_1, x_2, \dots, x_n with rational coefficients. Let $k \leq n$ be a positive integer and let $a_1 \geq a_2 \geq \cdots \geq a_k$ be positive integers. We denote by $\langle a_1, a_2, \dots, a_k \rangle$ the symmetric polynomial $\sum x_{i_1}^{a_1} x_{i_2}^{a_2} \cdots x_{i_k}^{a_k}$, where the sum is over all permutations of $\{1, 2, \dots, n\}$ that yield distinct monomials. Note that the degree of $\langle a_1, a_2, \dots, a_k \rangle$ is $a_1 + a_2 + \cdots + a_k$. For example, when $n = 3$, $\langle 1 \rangle = x_1 + x_2 + x_3$, $\langle 2 \rangle = x_1^2 + x_2^2 + x_3^2$, $\langle 1, 1 \rangle = x_1 x_2 + x_1 x_3 + x_2 x_3$, and $\langle 2, 1 \rangle = x_1^2 x_2 + x_1^2 x_3 + x_2^2 x_1 + x_2^2 x_3 + x_3^2 x_1 + x_3^2 x_2$. The notation $\langle a_1, a_2, \dots, a_k \rangle$ appears in [2] (p. 82, Ex. 5) and such a polynomial is called (see [1]) a monomial symmetric

function (and is also represented by m_λ corresponding to the partition $\lambda = (a_1, a_2, \dots, a_k)$). It will be important to determine products of monomial symmetric functions; note, for example, that with $n \geq 4$, $\langle 2 \rangle \langle 1, 1 \rangle = \langle 2, 1, 1 \rangle + \langle 3, 1 \rangle$ while $\langle 1, 1 \rangle \langle 1, 1 \rangle = 6\langle 1, 1, 1, 1 \rangle + 2\langle 2, 1, 1 \rangle + \langle 2, 2 \rangle$.

Theorem 1. *Let $m(1), m(2), \dots, m(n)$ be monomial symmetric functions in $Q[x_1, x_2, \dots, x_n]$, with $m(r)$ of degree r . Then $m(1), m(2), \dots, m(t)$ generate $Q[p_1, p_2, \dots, p_t]$, for $t = 1, 2, \dots, n$.*

We say that a set S in Σ is a *basis* of Σ if S generates Σ (i.e. $\Sigma = Q[S]$) and each element of Σ has a unique representation in $Q[S]$. It is well-known that the set of elementary symmetric functions forms a basis of Σ , as does $\{p_1, p_2, \dots, p_n\}$.

Theorem 2. *The set $\{m(1), m(2), \dots, m(n)\}$ defined above forms a basis of Σ .*

The proof makes use of the following lemma.

Lemma. *Consider the monomial symmetric function $\langle a_1, a_2, \dots, a_k \rangle$ in $Q[x_1, x_2, \dots, x_n]$ and let $t = \sum_{i=1}^k a_i$. Then there is a positive rational number c and an element B in $Q[p_1, p_2, \dots, p_{t-1}]$ such that*

$$\langle a_1, a_2, \dots, a_k \rangle = (-1)^{k-1} c p_t + B.$$

Proof: If $k = 1$ the lemma is true (since $\langle a_1 \rangle = p_1$). Assuming that the lemma is true when k is replaced by $k - 1$, we prove it for k .

Note that the product $\langle a_1, a_2, \dots, a_{k-1} \rangle \langle a_k \rangle$ is a symmetric polynomial and the term

$$x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}$$

appears a nonzero number of times in the expansion of the above product; thus

$$\langle a_1, a_2, \dots, a_{k-1} \rangle \langle a_k \rangle = d \langle a_1, \dots, a_k \rangle + \sum d_i A_i,$$

where d and each d_i is a positive integer and each A_i has the form $\langle b_1, b_2, \dots, b_{k-1} \rangle$. By the induction assumption

$$A_i = (-1)^{k-2} c_i p_t + D_i$$

where c_i is a positive rational number and D_i lies in $Q[p_1, p_2, \dots, p_{t-1}]$. Thus,

$$\langle a_1, a_2, \dots, a_{k-1} \rangle \langle a_k \rangle = d \langle a_1, a_2, \dots, a_k \rangle + \left(\sum_i (-1)^{k-2} d_i c_i \right) p_t + \sum d_i D_i.$$

Noting that since $\langle a_1, a_2, \dots, a_{k-1} \rangle$ and $\langle a_k \rangle$ have degrees less than t , they are in $Q[p_1, p_2, \dots, p_{t-1}]$, and thus so is their product, we complete the induction by solving the last equation for $\langle a_1, a_2, \dots, a_k \rangle$. \square

We are now ready to prove theorem 1.

Proof: The theorem is clearly true when $t = 1$, since $m(1) = p_1$. Assuming it is true when t is replaced by $t - 1$, we prove it for t .

Let P be in $Q[p_1, p_2, \dots, p_t]$. Then

$$P = \sum_{i=0}^r P_i p_t^i \tag{1}$$

where each P_i is in $Q[p_1, p_2, \dots, p_{t-1}]$, and thus by the induction hypothesis, in $Q[m(1), m(2), \dots, m(t-1)]$. By the lemma, $p_t = c' m(t) + B$ where c' is a nonzero rational number and B is in $Q[p_1, p_2, \dots, p_{t-1}]$ hence in $Q[m(1), m(2), \dots, m(t-1)]$. Replacing p_t in (1) by $c' m(t) + B$ completes the proof.

If a polynomial P in Σ had more than one representation in $Q[m(1), m(2), \dots, m(n)]$ then, from the lemma, P would have more than one representation in $Q[p_1, p_2, \dots, p_n]$, which is a contradiction. Thus every P in Σ can be represented uniquely as an element of $Q[m(1), m(2), \dots, m(n)]$ and hence $\{m(1), m(2), \dots, m(n)\}$ is a basis of Σ . However, there is a direct approach to establishing the uniqueness, suggested by Noah Brannen.

Consider the vector space V over Q of all symmetric polynomials of degree t with rational coefficients, together with the zero polynomial. Its dimension is the cardinality of the set

$$\{(i_1, i_2, \dots, i_n) : i_j \in \mathbb{Z}, 0 \leq i_1 \leq i_2 \leq \dots \leq i_n, \sum i_j = t\}.$$

We know that the set

$$W = \{(m(1))^{j_1} (m(2))^{j_2} \dots (m(t))^{j_t} : j_k \geq 0, \sum k j_k = t\}$$

generates V . But it is not difficult to show that the cardinality of W is equal to the dimension of V . Hence W is linearly independent, which is equivalent to the uniqueness of representation by $m(1), m(2), \dots, m(n)$.

Remark. Though we proved the theorem only when the coefficient field is Q , it holds for all fields of characteristic 0. It does not hold if the characteristic is two, since, for example, p_1 and p_2 do not generate the symmetric polynomials in x_1 and x_2 with coefficients in $GF(2)$ ($x_1 x_2$ is not an element of $GF(2)[p_1, p_2]$). However the theorem does hold if the characteristic of the coefficient field is larger than n . This follows from the fact that the rational number c in the lemma is $(k-1)! / \sum_{i=1}^k (r_i)!$ where r_i is the number of a_j equal to i (as can be proved by induction on k).

REFERENCES

1. I. G. MacDonald, *Symmetric Functions and Hall Polynomials*, Southampton: The Camelot Press Ltd., 1979.
2. B. L. van der Waerden, *Modern Algebra*, Vol 1, second ed., N.Y., Frederick Ungar Publishing Co., 1943.

Mathematics Department
University of California at Davis
Davis, CA 95616-8633

A Formula and a Proof of the Infinitude of the Primes

Michael Rubinstein

In this short note we present a proof that there exist infinitely many primes. The basic idea runs as follows. Consider a set of n distinct primes $S = \{p_1, p_2, p_3, \dots, p_n\}$. We ask: how many positive integers $\leq x$ are generated by S

(i.e. are of the form $p_1^{\beta_1} p_2^{\beta_2} \dots p_n^{\beta_n}$ where the β_i 's are non-negative integers)? We obtain an asymptotic answer to this question from which we deduce that there cannot be finitely many primes.

The Formula. Let $S = \{p_1, p_2, p_3, \dots, p_n\}$ where the p_i 's are distinct primes. Let $f(S, x)$ denote the number of positive integers $\leq x$ that are of the form $p_1^{\beta_1} p_2^{\beta_2} \dots p_n^{\beta_n}$ where the β_i 's are non-negative integers. Then,

$$f(S, x) \sim \frac{(\log x)^n}{n! \prod_{i=1}^n \log p_i}. \quad (1)$$

Note that (1) would imply the existence of infinitely many primes. For, say we only had finitely many primes. Denote them by $S = \{p_1, p_2, p_3, \dots, p_n\}$. Then, S would have to generate all the integers, i.e. $f(S, x)$ would have to be equal to $[x]$ (where $[x]$ is the floor function). But, this would imply that

$$[x] \sim \frac{(\log x)^n}{n! \prod_{i=1}^n \log p_i}$$

which is a contradiction (since $\lim_{x \rightarrow \infty} (\log x)^n / [x] = 0$).

Actually, we do not need an asymptotic result to deduce that there are infinitely many primes. For example, it is easy to see that $f(S, x) \leq ([\log_2 x] + 1)^n$ from which the infinitude of the primes follows (this argument is essentially the same as an argument presented by Thue. See [1] for his argument as well as several other proofs of the infinitude of the primes). However, the beauty of the asymptotic formula encourages us to continue.

Derivation of (1). We prove (1) by establishing the following two inequalities:

$$f(S, x) \geq \frac{(\log x)^n}{n! \prod_{i=1}^n \log p_i} \quad (A)$$

$$f(S, x) \leq \frac{(\log x)^n}{n!} \mu_n + \frac{(\log x)^{n-1}}{(n-1)!} \mu_{n-1} + \dots + \frac{(\log x)}{1!} \mu_1 + 1, \quad (B)$$

where

$$\begin{aligned} \mu_1 &= \frac{1}{\log p_1} + \frac{1}{\log p_2} + \dots + \frac{1}{\log p_n} \\ \mu_2 &= \frac{1}{\log p_1 \log p_2} + \frac{1}{\log p_1 \log p_3} + \dots + \frac{1}{\log p_{n-1} \log p_n} \\ &\vdots \\ \mu_n &= \frac{1}{\log p_1 \log p_2 \dots \log p_n}, \end{aligned}$$

i.e., μ_1, \dots, μ_n are the elementary symmetric functions of $(1/\log p_1), \dots, (1/\log p_n)$. It is clear that (A) and (B), once proven, would imply (1).

Proof (by induction) of (A) and (B).

First of all $f(\{p_1\}, x)$ is equal to $[\log_{p_1} x] + 1$, where $[]$ is the floor function, and so satisfies (A) and (B) (since $(\log x / \log p_1) < [\log_{p_1} x] + 1 \leq (\log x / \log p_1) + 1$).

Next, assume that (A) and (B) hold for n elements, and consider the $n + 1$ case (i.e. the case where $S = \{p_1, p_2, p_3, \dots, p_{n+1}\}$). Now,

$$f(\{p_1, p_2, p_3, \dots, p_{n+1}\}, x) = \sum_{j=0}^{[\log_{p_{n+1}} x]} f\left(\{p_1, p_2, p_3, \dots, p_n\}, \frac{x}{(p_{n+1})^j}\right). \quad (2)$$

This can be shown by counting how many of the $f(\{p_1, p_2, p_3, \dots, p_{n+1}\}, x)$ numbers contain no powers of p_{n+1} (corresponds to $j = 0$), contain one power of p_{n+1} (corresponds to $j = 1$), contain two powers of p_{n+1} (corresponds to $j = 2$) etc.

By (A) (and by our induction hypothesis), (2) is greater than or equal to

$$\sum_{j=0}^{[\log_{p_{n+1}} x]} \left(\frac{\left(\log \frac{x}{(p_{n+1})^j} \right)^n}{n! \prod_{i=1}^n \log p_i} \right), \quad (3)$$

and, by (B) (as well as our induction hypothesis), (2) is less than or equal to

$$\begin{aligned} & \sum_{j=0}^{[\log_{p_{n+1}} x]} \frac{\left(\log \frac{x}{(p_{n+1})^j} \right)^n}{n!} \mu_n + \frac{\left(\log \frac{x}{(p_{n+1})^j} \right)^{n-1}}{(n-1)!} \mu_{n-1} \\ & + \dots + \frac{\left(\log \frac{x}{(p_{n+1})^j} \right)}{1!} \mu_1 + 1. \end{aligned} \quad (4)$$

Applying Lemma 1 (which we prove shortly—look at the Lemma) we have that (3) (and thus (2)) is greater than or equal to

$$\frac{(\log x)^{n+1}}{(n+1)! \prod_{i=1}^{n+1} \log p_i},$$

and that (4) (and thus (2)) is less than or equal to

$$\frac{(\log x)^{n+1}}{(n+1)!} \sigma_{n+1} + \frac{(\log x)^n}{n!} \sigma_n + \dots + \frac{(\log x)}{1!} \sigma_1 + 1,$$

where $\sigma_1, \dots, \sigma_{n+1}$ are the elementary symmetric functions of $(1/\log p_1), \dots, (1/\log p_{n+1})$ (to prove this, use $\sigma_i = \mu_i + (1/\log p_{n+1})\mu_{i-1}$, $i = 1, 2, \dots, n+1$ (where we set $\mu_0 = 1$ and $\mu_{n+1} = 0$)).

Thus, (A) and (B) have been established and, therefore, (1) follows.

One property was used in the above which we list and prove in the following lemma.

Lemma 1.

$$\frac{(\log x)^{n+1}}{(n+1)\log k} \leq \sum_{j=0}^{\lfloor \log_k x \rfloor} \left(\log \frac{x}{k^j} \right)^n \leq (\log x)^n + \frac{(\log x)^{n+1}}{(n+1)\log k},$$

where $k > 1$, $n \geq 0$, and $x \geq k$.

Proof.

We compare the sum to $\int (\log(x/k^t))^n dt$. On the interval $(-\infty, \log_k x]$, the function $g(t) = (\log(x/k^t))^n$ is a monotonically decreasing non-negative function (which is easy to see. For those that cannot see this, differentiating the function will also work). Thus,

$$\int_0^{\log_k x} \left(\log \frac{x}{k^t} \right)^n dt \leq \sum_{j=0}^{\lfloor \log_k x \rfloor} \left(\log \frac{x}{k^j} \right)^n$$

(note: when $t > \log_k x$, $\log(x/k^t)$ is negative. Thus, to prevent running into trouble, for example, when $n = 1/2$, and also to obtain a stronger inequality (when n is an odd integer), we integrate from 0 to $\log_k x$ rather than from 0 to $\lfloor \log_k x \rfloor + 1$).

Substituting $w = (x/k^t)$ and then $u = \log w$ into the integral, we obtain the left hand side of the Lemma.

The right hand side of the Lemma may be obtained in a similar fashion:

$$\begin{aligned} \sum_{j=0}^{\lfloor \log_k x \rfloor} \left(\log \frac{x}{k^j} \right)^n &= (\log x)^n + \sum_{j=1}^{\lfloor \log_k x \rfloor} \left(\log \frac{x}{k^j} \right)^n \\ &\leq (\log x)^n + \int_0^{\lfloor \log_k x \rfloor} \left(\log \frac{x}{k^t} \right)^n dt \leq (\log x)^n + \int_0^{\log_k x} \left(\log \frac{x}{k^t} \right)^n dt \\ &= (\log x)^n + \frac{(\log x)^{n+1}}{(n+1)\log k}. \end{aligned}$$

The last step can be shown using the substitutions recommended earlier.

An Alternative Approach. We can also obtain (1) in the following manner. With each $p_1^{\beta_1} p_2^{\beta_2} \dots p_n^{\beta_n}$ we associate the n -dimensional point $(\beta_1, \beta_2, \dots, \beta_n)$. We are interested in counting the number of non-negative integer solutions $(\beta_1, \beta_2, \dots, \beta_n)$ to

$$1 \leq p_1^{\beta_1} p_2^{\beta_2} \dots p_n^{\beta_n} \leq x \quad (5)$$

Taking logarithms this becomes

$$0 \leq \sum_{i=1}^n \beta_i \log p_i \leq \log x$$

from which we see that the number of solutions to (5) is equal to the number of integer points contained in and on the n -dimensional polyhedron with vertices at $(0, 0, \dots, 0)$, $((\log x / \log p_1), 0, \dots, 0)$, \dots , $(0, 0, \dots, (\log x / \log p_n))$. Thus, we see that the number of solutions to (5) is equal to the volume of this polyhedron + $O(\text{surface area of this polyhedron})$. The volume of this polyhedron is evaluated (say by integration) and is found to equal

$$\frac{(\log x)^n}{n! \prod_{i=1}^n \log p_i}$$

while the surface area is found to equal

$$\frac{(\log x)^{n-1} \sum_{i=1}^n \log p_i}{(n-1)! \prod_{i=1}^n \log p_i} + \frac{(\log x)^{n-1} \sqrt{(\log p_1)^2 + \dots + (\log p_n)^2}}{(n-1)! \prod_{i=1}^n \log p_i} = O((\log x)^{n-1})$$

from which formula (1) follows.

Conclusion. As a final observation, note that this whole exposition still works if, instead of prime numbers, we let S consist of integers > 1 that are relatively prime in pairs. For, in so doing, unique factorization still holds. Our methods only work if two distinct tuples $(\alpha_1, \alpha_2, \dots, \alpha_n)$ and $(\beta_1, \beta_2, \dots, \beta_n)$ give rise to two distinct integers.

REFERENCE

- [1] Paulo Ribenboim, *The Book of Prime Number Records*, Springer Verlag, New York, 1988, pg 8.

Princeton University
Princeton, NJ 08544
miker@phoenix.princeton.edu

The Risks of Self-Evaluation

Writing of his recent "discovery of a vast theory of double determinants" (in present day terms, determinants of matrices whose terms are determinants), J. J. Sylvester described it as "the dawn of a new epoch in the history of modern algebra." (*Philosophical Magazine*, 25 (1863) p. 453; *Collected Mathematical Papers*, vol. 2, p. 331.) He wrote no further on the subject!

Kenneth O. May

—*American Mathematical Monthly*
72, (1965) p. 314.

UNSOLVED PROBLEMS

Edited by: **Richard Guy**

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

Are There Only Finitely Many Binomial Coefficients With Positive Deficiency?

The product of the k consecutive numbers

$$n + 1, n + 2, \dots, n + i, \dots, n + k$$

is, of course, divisible by $k!$; that is, the binomial coefficient $\binom{n+k}{k}$ is an integer.

For each i , $1 \leq i \leq k$, write the number $n + i$ as a product $a_i b_i$ where a_i contains those prime factors which are at most k , and b_i contains those greater than k . Then $\prod a_i$ is a multiple of $k!$.

If $\prod a_i$ is exactly equal to $k!$, then Erdős, Lacampagne & Selfridge [2] describe the binomial coefficient $\binom{n+k}{k}$ as being **good**.

$$\binom{23}{5} = 7 \times 11 \times 19 \times 23$$

has no prime factor less than or equal to 5, and so is good.

In an earlier paper [1] they define the **deficiency** of a good binomial coefficient as the number of b_i which are equal to 1. The deficiency of $\binom{23}{5}$ is thus one (just $b_2 = 1$), while

$$\begin{aligned} \binom{284}{28} &= 257 \cdot 43 \cdot 37 \cdot 29 \cdot 131 \cdot 263 \cdot 53 \cdot 89 \cdot 67 \cdot 269 \cdot 271 \cdot 137 \cdot 277 \\ &\quad \cdot 139 \cdot 31 \cdot 281 \cdot 47 \cdot 283 \cdot 71 \end{aligned}$$

with $b_i = 1$ for $i = 4, 8, 10, 14, 16, 17, 19, 20$ and 24, has deficiency nine, the largest known.

The authors of [1 & 2] believe that they have found all binomial coefficients with deficiency at least two, namely $\binom{284}{28}$; $\binom{47}{11}$ with deficiency 4; $\binom{46}{10}$, $\binom{47}{10}$, $\binom{241}{16}$,

$\binom{2105}{25}$, $\binom{1119}{27}$ and $\binom{6459}{33}$, each of deficiency 3; and $\binom{7}{4}$, $\binom{44}{8}$, $\binom{74}{10}$, $\binom{174}{12}$, $\binom{239}{14}$, $\binom{5179}{27}$, $\binom{8413}{28}$, $\binom{8414}{28}$ and $\binom{96622}{42}$, with deficiencies 2.

They also believe that

¿ there are only finitely many binomial coefficients with deficiency one ?

and they further conjecture that the least prime factor of the binomial coefficient $\binom{N}{k}$ is less than or equal to the maximum of N/k and 29, these maxima being attained in the examples $\binom{215}{5}$ and $\binom{284}{28}$ mentioned above.

The truth of the famous Conjecture H of Schinzel [4, 5] would imply that there are infinitely many good binomial coefficients for which each of the b_i is prime. These have deficiency zero and are the product of k primes. For example,

$$\binom{215}{5} = 43 \times 53 \times 71 \times 107 \times 211.$$

Erdős & Selfridge [3] noted that if $n \geq 2k \geq 4$, then there is at least one value of i , $0 \leq i \leq k-1$, such that $n-i$ does not divide $\binom{n}{k}$, and asked for the least n_k for which there was only one such i . For example, $n_2 = 4$, $n_3 = 6$, $n_4 = 9$, $n_5 = 12$. $n_k \leq k!$ for $k \geq 3$.

REFERENCES

1. P. Erdős, C. B. Lacampagne & J. L. Selfridge, Prime factors of binomial coefficients and related problems, *Acta Arith.*, 49 (1988) 507–523.
2. P. Erdős, C. B. Lacampagne & J. L. Selfridge, Estimates on the least prime factor of a binomial coefficient, *Math. Comput.*, 60 (1993).
3. P. Erdős & J. L. Selfridge, Problem 6447, this MONTHLY, 90 (1983) 710; 92 (1985) 435–436.
4. A. Schinzel & W. Sierpiński, Sur certaines hypothèses concernant les nombres premiers, *Acta Arith.*, 4 (1958) 185–208; corrigendum 5 (1960) 259.
5. W. Sierpiński, *Elementary Theory of Numbers*, 2nd English edition (ed. A. Schinzel), PWN Warszawa & Elsevier New York, 1987, Ch. 3, §8, p.133.

I hear and I forget. I see and I remember. I do and I understand.

—Chinese Proverb

PROBLEMS AND SOLUTIONS

Edited by:

Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before September 30, 1993 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10298. *Proposed by Donald E. Knuth, Stanford University, Stanford, CA.*

Let $\left\{ \begin{smallmatrix} m+n \\ n \end{smallmatrix} \right\}$ denote the number of ways to partition a set of $m+n$ elements into n nonempty subsets. Prove that

$$\frac{2^m 3^{\lfloor m/2 \rfloor} 4^{\lfloor m/3 \rfloor} 5^{\lfloor m/4 \rfloor} \cdots}{(n+1)(n+2) \cdots (n+m)} \left\{ \begin{smallmatrix} m+n \\ n \end{smallmatrix} \right\}$$

is an integer.

10299. *Proposed by José Luis Palacios, New Jersey Institute of Technology, Newark, NJ.*

For an odd integer N , greater than 3, let G_N be the cyclic graph on N vertices. Consider the following random walk of *two* particles on G_N : at each time step, both particles independently move to one of the two adjacent vertices with probability $1/2$.

If the particles initially occupy adjacent vertices, what is the expected number of jumps until the particles meet?

10300. Proposed by Eric H. Mason, Herndon, VA.

Let $\{P_m(z): m = 1, 2, 3, \dots\}$ be the sequence of polynomials defined by $P_1(z) = z - 1$, $P_2(z) = z^2 - z - 1$, and $P_m(z) = zP_{m-1}(z) - P_{m-2}(z)$ for $m > 2$. Show that the roots of $P_m(z)$ are $2 \cos((2k - 1)\pi/(2m + 1))$ for $1 \leq k \leq m$.

10301. Proposed by William P. Wardlaw, United States Naval Academy, Annapolis, MD.

Let R be a commutative ring with identity. For which matrices A in $\mathbf{GL}_n(R)$ is the mapping

$$\alpha_A: \mathbf{SL}_n(R) \rightarrow \mathbf{SL}_n(R) \text{ defined by } X \mapsto AXA^{-1}$$

an inner automorphism of $\mathbf{SL}_n(R)$.

10302. Proposed by Jeffrey A. Barnett, Northrop Corporation, Palos Verdes Peninsula, CA.

Let \mathcal{J} be the class of integer-valued functions defined on \mathbb{N}^+ that satisfy the following constraints. For $j \in \mathcal{J}$, $j(1) = 1$ and

$$j(m) + j(n) - 1 \leq j(m + n) \leq j(m) + j(n)$$

for all $m, n \in \mathbb{N}^+$. Show that the number of distinct initial sequences of length N generated by the $j \in \mathcal{J}$ is

$$\sum_{1 \leq n \leq N} \phi(n)$$

where $\phi(n)$ is Euler's totient function.

10303. Proposed by David E. Gurarie, Case Western Reserve University, Cleveland, OH.

Let a_1, \dots, a_n ($n \geq 3$) be positive real numbers.

(a) Find necessary and sufficient conditions on a_1, \dots, a_n for there to exist a convex n -gon which admits an inscribed circle and whose sides, in cyclic order, are a_1, \dots, a_n .

(b) Find the radius of the inscribed circle.

10304. Proposed by Ignacy I. Kotlarski, Oklahoma State University, Stillwater, OK.

Let $\lambda_0, \lambda_1, \lambda_2$ be three positive constants. Let X_0, X_1, X_2 be three independent discrete random variables with nonnegative integer values only. Suppose that $EX_0 = \lambda_0$. Now let $Y_1 = X_0 + X_1$ and $Y_2 = X_0 + X_2$ and suppose that the joint probability distribution for (Y_1, Y_2) is given by

$$P(Y_1 = j_1, Y_2 = j_2) = \sum_{k=0}^{\min(j_1, j_2)} \frac{\lambda_0^k \lambda_1^{j_1-k} \lambda_2^{j_2-k}}{k!(j_1-k)!(j_2-k)!} e^{-(\lambda_0 + \lambda_1 + \lambda_2)}$$

for nonnegative integers j_1 and j_2 . Find the distribution of X_0, X_1 , and X_2 .

10305. *Proposed by J. C. Lagarias, AT&T Bell Laboratories, Murray Hill, NJ.*

Is there a smallest prime number if the set of primes is enlarged to include certain real algebraic numbers? In particular:

(a) Call a real algebraic number α an A -prime-number if it is a totally real algebraic integer such that:

(i) The ideal (α) is a prime ideal in the ring of integers of the field $\mathbb{Q}(\alpha)$.

(ii) $\alpha \geq |\sigma(\alpha)|$ for all algebraic conjugates $\sigma(\alpha)$ of α .

Is there a smallest A -prime-number?

(b) Call a number of A^* -prime-number if it is a real algebraic integer (not necessarily totally real) satisfying (i) and (ii) above.

Is there a smallest A^* -prime-number?

NOTES

(10301) $\mathbf{GL}_n(R)$ denotes the group, under matrix multiplication, of all invertible n by n matrices and $\mathbf{SL}_n(R)$ is the subgroup of $\mathbf{GL}_n(R)$ of matrices with determinant 1. An *inner automorphism* of a group is one given by the formula for α_A with A in the group. (10302) Here \mathbb{N}^+ is used to denote the positive integers, and $\phi(n)$ is the number of elements of \mathbb{N}^+ less than or equal to n that are relatively prime to n . (10303) The cases $n = 3$ and $n = 4$ are well known: for $n = 3$, the only condition is that the sides form a triangle, and the inscribed circle is uniquely determined by the sides; for $n = 4$, a necessary condition for the existence of an inscribed circle is $a_1 + a_3 = a_2 + a_4$, and there is a continuum of circles which can be inscribed in quadrilaterals having these sides. An answer to part b should include a suitable *formulation* in this case. (10304) The given distribution of (Y_1, Y_2) is the two dimensional Poisson distribution with parameters $\lambda_0, \lambda_1, \lambda_2$. (10305) The A -prime-numbers in \mathbb{Z} are exactly the primes. The *algebraic conjugates* are all roots of a polynomial with rational coefficients irreducible over the rationals. An algebraic integer α is *totally real* if all algebraic conjugates $\sigma(\alpha)$ are real.

SOLUTIONS

Invertible Derivatives of Rational Functions

6656 [1991, 372]. *Proposed by Dragomir Ž. Đoković, University of Waterloo, Ontario.*

(i) If $P(z)$ is a non-zero polynomial over \mathbb{C} with at least two distinct zeros, prove that $1/P(x)$ has at least one non-zero residue (or equivalently $1/P(z)$ is not the derivative of a rational function).

(ii) Prove or disprove the following more general assertion: If f and g are rational functions over \mathbb{C} such that $f'(z)g'(z) = 1$, then $f'(z) = c(z - a)^k$ for some complex constants c, a and some integer $k \neq \pm 1$.

Solution part (i) by Robin J. Chapman, University of Exeter, Exeter, U. K. Let $P(z)$ have degree d and suppose it has zeros a_1, a_2, \dots, a_m of orders d_1, d_2, \dots, d_m respectively. By hypothesis $m \geq 2$ and each $d_i \geq 1$. Now if $1/P(z) = Q'(z)$ with $Q(z)$ a rational function then clearly $Q(z)$ has poles of orders $d_i - 1$ at the a_i but at no other points. Hence $Q(z) = f(z)/g(z)$ where $f(z)$ and $g(z)$ are polynomials of degrees r and $s = d - m$ respectively. Without loss of generality $r \neq s$ as otherwise we can subtract a constant from $Q(z)$ to make this so. Expand $[P(z)]^{-1}$ and $Q(z)$ as Laurent series in $1/z$. We have

$$1/P(z) = \sum_{i=d}^{\infty} b_i z^{-i}$$

and

$$Q(z) = \sum_{j=s-r}^{\infty} c_j z^{-j}$$

where $b_d \neq 0$ and $c_{s-r} \neq 0$. Differentiating we get

$$Q'(z) = - \sum_{j=s-r}^{\infty} j c_j z^{-j-1}$$

which can only equal $[P(z)]^{-1}$ if $d = s - r + 1$. So $d = s - r + 1 \leq s + 1 = d - m + 1$ which is impossible as $m \geq 2$.

Solution of part (ii) by F. J. Flanigan, San Jose State University, San Jose, CA. Let $f'(z) = rz^{3r-1}(a + bz^r)^{-n}$ when $n \geq 4$ and $r \geq 3$, both integers. This f' arises from substituting $x = z^r$ in the integral $\int x^2(a + bx)^{-n} dx$ which, for $n \neq 1, 2$, or 3 , works out to a rational function given explicitly in integral tables, say, *CRC Handbook*, 10th edition, 1956, p. 270, #39. The reciprocal $1/f'(z) = (a + bz^r)^n / (rz^{3r-1})$ has a *terminating* Laurent expansion at $z = 0$

$$b_0 z^{1-3r} + b_1 z^{1-2r} + b_2 z^{1-r} + b_3 z + \dots$$

If $r \neq 1$ or 2 , then this expression obviously has a rational antiderivative.

So setting $g'(z) = 1/f'(z)$ yields a family of pairs f and g of rational functions which are counterexamples to the assertion of part ii).

Editorial comment. The two conditions in part (i) are equivalent, for if a meromorphic function $h(z)$ has no non-zero residue, then a single-value primitive $H(z)$ is constructed by integrating along any path (not passing through a pole) from a fixed base point to z . A trivial Laurent series argument shows that only pole singularities will arise.

F. J. Flanigan noted that part (i) has already appeared in this MONTHLY as problem E2236 [1970, 522] with an erroneous solution [1971, 408] and a correct solution [1971, 905].

Yet another approach to part (i) was given by an editorial consultant who observed that if $1/P = F'$ for some rational F , then we may assume that $F(\infty) = 0$ and write $F = 1/f$ where the rational function f has a pole at infinity. Counting zeros and poles of f and f' and noting that every finite zero of f' must also be a zero of f leads to the conclusion that either f or $1/f$ is a polynomial with only one zero (of some multiplicity).

This same consultant also discovered a family of counterexamples to part (ii) in which $f' = (P/Q)^2$ where P and Q are polynomials satisfying $(PQ)'' = 4P'Q'$. An example similar to those given above has $P(z) = z^{l(l-1)/2}$ and $Q(z) = z^{l(l+1)/2} + Az^{(l-1)(l-2)/2}$ which gives $f(z) = (z^{2l-1} + A)^{-1}$. A further example has $P(z) = z^3 + z^2 + z/3$ and $Q(z) = z^6 + 2z^5 + 5z^4/3 + 5z^3/9$ which leads to

$$f(z) = -\frac{3}{5} \frac{3z+1}{z^3(9z^3+18z^2+15z+5)}$$

$$g(z) = \frac{z^7}{7} + \frac{z^6}{3} + \frac{z^5}{3} + \frac{z^4}{9} - \frac{z^3}{27} + \frac{z}{81} + \frac{1+2z}{81(1+3z+3z^2)}.$$

Solved by the two respondents and one editorial consultant cited above. One incorrect solution was also received.

Extraneous Primes

E 3452 [1991, 645]. Proposed by C. A. Nicol and J. L. Selfridge, University of South Carolina, Columbia, SC.

If n is an odd integer greater than 3 and ϕ is the Euler function, prove that there exists a prime p such that $p|(2^{\phi(n)} - 1)$ but $p \nmid n$.

Solution I by David Callan, University of Wisconsin, Whitewater, WI. Suppose n has k distinct prime factors. Then $\phi(n)$ is divisible by 2^k , and so $2^{\phi(n)} - 1 = Y^{2^k} - 1$, with Y an even integer greater than 3. But $Y^{2^k} - 1 = (Y - 1)\prod_{i=0}^{k-1}(Y^{2^i} + 1)$ is a product of $k + 1$ factors that are pairwise relatively prime. Hence $2^{\phi(n)} - 1$ has more than k prime factors and the result follows.

Solution II by Kenneth Rogers, Harvard University, Cambridge, MA. If n has at least two distinct odd prime factors and p is one of them, then $(p - 1)|(\phi(n)/2)$, and hence $p|(2^{\phi(n)/2} - 1)$. Since $2^{\phi(n)} - 1 = (2^{\phi(n)/2} - 1)(2^{\phi(n)/2} + 1)$ and the two factors are relatively prime, we have $(2^{\phi(n)/2} + 1, n) = 1$; hence every prime divisor of $2^{\phi(n)/2} + 1$ is prime to n but divides $2^{\phi(n)} - 1$. This leaves only the cases when $n = p^m$. If $n = 3^m$ with $m \geq 2$, then $3|\phi(n)$, and hence 7 divides $2^{\phi(n)} - 1$ but not n . If $n = p^m$ with $p > 3$, then $2|\phi(n)$, and hence 3 divides $2^{\phi(n)} - 1$ but not n .

Editorial comment. Volker Strehl and P. G. Walsh independently showed that the result of the problem holds for all n except $\{1, 2, 3, 6\}$. D. W. Kosler, Gerry Myerson and Hongbing Yu each went further and investigated the set of pairs a, n such that $(a, n) = 1$, $a > 1$ and $n > 2$ for which there is a prime p such that $p|(a^{\phi(n)} - 1)$ but $p \nmid n$. The complete list of such (n, a) is $\{(3, 2), (4, 3), (6, 5), (6, 7), (6, 17), (10, 3)\}$. The editors thank N. J. Fine for suggesting the title.

Solved by 46 readers (including those cited) and the proposers. One incomplete and two incorrect solutions were received.

Pigeons on a Circle

E 3453 [1991, 645]. Proposed by Allen J. Schwenk, Western Michigan University, Kalamazoo, MI.

Suppose n is a positive integer greater than 2. Determine the smallest positive number c_n with the following property: Given any n distinct real numbers, there

exist two of them, say x and y , which satisfy

$$0 < \frac{x - y}{1 + xy} \leq c_n.$$

Solution by Ilias Kastanas, California State University, Los Angeles, CA. The answer is $c_n = \tan(\pi/n)$.

For $x_1 < x_2 < \cdots < x_n$, let $x_i = \tan(\theta_i)$ with $-\pi/2 < \theta_i < \pi/2$. Then $(x_i - x_j)/(1 + x_i x_j) = \tan(\theta_i - \theta_j)$. Since the n positive numbers $\theta_2 - \theta_1, \theta_3 - \theta_2, \dots, \theta_n - \theta_{n-1}, \pi + \theta_1 - \theta_n$ add up to π , at least one of them is less than or equal to π/n , and so its tangent is less than or equal to $\tan(\pi/n)$.

The choice $\theta_2 - \theta_1 = \theta_3 - \theta_2 = \cdots = \theta_n - \theta_{n-1} = \pi/n$ (say with $\theta_1 = -\pi/2 + \pi/(2n)$) shows that this value of c_n is best possible.

Editorial comment. The problem generalizes and optimally tightens the inequality of MONTHLY Problem E3121 [1985, 736], in which one was asked to prove that $c_n \leq \tan(\pi/(n-1))$ for $n = 13$. The large number of incorrect solutions to the present problem was due to overlooking the possibility that $(x_1 - x_n)/(1 + x_1 x_n)$ could be small, and hence making an assumption that the result of E3121 was optimal. In a successful proof, one applies the *pigeonhole principle* on the circle. This observation led Kevin Ford to suggest the title for this solution.

Solved also by R. J. Chapman (U. K.), P. Čížek (student, Czech Republic), M. Dindos (Slovakia), K. Ford (student), V. Glasnák (student, Czech Republic), I. Goldberg (Canada), L. Guijarro & S. E. Arteaga, R. Holzsager, N. Komanda, K.-W. Lau (Hong Kong), O. P. Lossers (The Netherlands), D. Magagnosc, H. Morris, A. Müller (France), J. H. Nieto (Venezuela), A. Pedersen (Denmark), S. G. Penrice, M. Roth & O. Šuch (Canada), D. L. Stock, A. Swett, D. B. Tyler, D. Velleman, B. M. M. de Weger (The Netherlands), K. Zacharias (Germany), Anchorage Math Solutions Group, National Security Agency Problems Group, and the proposer. Thirteen incorrect solutions were received.

Collecting Time on The Cyclic Graph

6665 [1991, 655]. *Proposed by José Luis Palacios, New Jersey Institute of Technology, Newark, NJ, and Dennis P. Sandell, Swedish University of Agricultural Sciences, Garpenberg, Sweden.*

Let S_n ($n \geq 0$) be a simple symmetric random walk, i.e., $S_0 = 0$ and $S_n = X_1 + X_2 + \cdots + X_n$ for $n > 0$, where the X_i are independent identically distributed random variables with $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$. Let N be an arbitrary positive integer and let T be the first time that the difference between the maximum and minimum of the random walks is N , i.e., let

$$T = \min \left\{ n : \max_{0 \leq k \leq n} S_k - \min_{0 \leq k \leq n} S_k = N \right\}.$$

Find the expected value of T .

Solution I by Robin J. Chapman, University of Exeter, Exeter, U.K. The expected value of T_N is $1 + 2 + \cdots + N = N(N+1)/2$.

Let $U_1 = T_1$ and $U_N = T_N - T_{N-1}$ for $N \geq 2$. I shall prove that $E(U_N) = N$ from which the result is immediate.

It is obvious that $U_1 = 1$ with probability 1, so suppose $N \geq 2$. If $a = \min_{0 \leq k \leq T_{N-1}} S_k$ and $b = \max_{0 \leq k \leq T_{N-1}} S_k$ then $b - a = n - 1$ and either $S_{T_{N-1}} = a$ or $S_{T_{N-1}} = b$. Now U_N is the time it subsequently takes for the random walk to

reach either $a - 1$ or $b + 1$. But this is the classical problem of a random walk with absorbing barriers, and so $E(U_N) = N$ (see G. Grimmett & D. Welch, *Probability: an Introduction*, Oxford, 1986, Theorem 10D, p. 161) as claimed.

Solution II by Donald A. Darling, Newport Beach, CA. There is little extra difficulty in treating the asymmetric walk in which $S_{n+1} = S_n \pm 1$ with probabilities $\frac{1}{2} \pm \frac{1}{2}\delta$, $|\delta| < 1$. In fact we can find the distribution (in the form of a generating function) of T . Let N_c be the least k , if any, such that $S_k = c$, and let $\phi_+ = \phi_+(t, \delta)$ be the generating function for the random variable N_1 , and ϕ_- be that of N_{-1} , i.e., $\phi_{\pm} = E(t^{N_{\pm 1}})$. It is well known and easy to prove that

$$\phi_{\pm} = \frac{1 - \sqrt{1 - (1 - \delta^2)t^2}}{(1 \mp \delta)t},$$

(cf. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, Ch. 14). If $c \geq 0$ the generating function of N_c is ϕ_+^c and that of N_{-c} is ϕ_-^c . The range R_n is defined as $\max_{0 \leq k \leq n} S_k - \min_{0 \leq k \leq n} S_k$ and T_r is the least integer n such that $R_n = r$.

In the following a and b are non-negative integers. Let $A_n^+ = A_n^+(a, b)$ be the event that $\{N_{-b} > n, N_a = n\}$, and similarly for $A_n^- = \{N_a > n, N_{-b} = n\}$. The event $\{N_a = n\}$ occurs if either A_n^+ occurs or if for some $v = 0, 1, \dots, n$ an event $A_v^- \cap \{N_{a+b} = n - v\}$ occurs. An analogous relation holds for the event $\{N_{-b} = n\}$. Letting $\psi_{\pm} = \psi_{\pm}(a, b) = \psi_{\pm}(a, b, t, \delta)$ be the generating functions for the probabilities of the events A_n^{\pm} we obtain

$$\phi_+^a = \psi_+ + \psi_- \phi_+^{a+b}$$

$$\phi_-^b = \psi_- + \psi_+ \phi_-^{a+b}$$

Setting $D = 1 - (\phi_+ \phi_-)^{a+b}$ these equations yield

$$\psi_+ = \frac{\phi_+^a (1 - (\phi_+ \phi_-)^b)}{D} \quad \psi_- = \frac{\phi_-^b (1 - (\phi_+ \phi_-)^a)}{D}$$

The union of the events $A_v^-(a, b) \cap A_{n-v}^+(a+b, 1)$ for $v = 1, 2, \dots, n$ is the event that $R_n = a + b$ and $T_a = n$, and has a generating function $\psi_-(a, b) \psi_+(a+b, 1)$. Similarly the event $R_n = a + b$ and $T_{-b} = n$ has the generating function $\psi_+(a, b) \psi_-(1, a+b)$. We add these two generating functions, express the sum in terms of the functions ϕ_{\pm} , and sum over non-negative values of a and b such that $a + b = r$. The result is the generating function for T_r . This is tedious but straightforward. The result can be put in a succinct form by introducing the function

$$g(x) = \frac{(1 - x^{r+1})(1 - x^r)}{1 - x}$$

with $g(1) = 0$. Then

$$f(t) = f(t, r, \delta) = E(t^{T_r}) = \frac{\phi_+^r g(\phi_-) + \phi_-^r g(\phi_+)}{g(\phi_+ \phi_-)}$$

and the mean is given by $f'(1)$. Setting $p = (1 - |\delta|)/(1 + |\delta|)$ further routine

calculation yields

$$f'(1) = E(T_r) = \frac{1}{|\delta|} \left(r - \frac{pg'(p)}{g(p)} + \frac{p^r g'(1)}{g(p)} \right)$$

and in the symmetric random walk, $\delta \rightarrow 0$, gives $E(T_r) = r(r+1)/2$.

Editorial comment. Quite a few readers noted that the problem as stated is equivalent to finding the mean time for a standard symmetric walk on the vertices of an $(n+1)$ -gon to have visited all the vertices, apart from the initial vertex; this is solved in H. Wilf, "The Editor's Corner: The White Screen Problem", this MONTHLY, 96 (1989), 704–707.

Solved also by R. A. Agnew, M. H. Andreoli, D. Beckwith, D. Callan, L. Crone & R. Holzinger, C. P. Grant, R. High, A. A. Jagers (The Netherlands), B. R. Johnson (Canada), I. Kastanas, K. S. Kedlaya (student), H. Morris, D. M. Rosenblum, M. Roth & O. Šuch (Canada), T. W. Starbird, C. Voas, E. A. Weinstein, D. Wolfe, and the proposers.

A Consequence of a Factorial Equation

6669 [1991, 767]. Proposed by Paul Erdős, Hungarian Academy of Sciences, Budapest, Hungary.

Prove that there is a constant c such that if $n! = a!b!$ with $1 < a < b < n$, then

$$n - b < c \log \log n.$$

(It has been conjectured, but never proved, that if $n > 10$ and $n! = a!b!$ with $1 < a < b < n$, then $b = n - 1$ and $n = a!$. For $n = 10$ there is the exception $10! = 6!7!.$)

Solution by the proposer. Let $n! = a!b!$ with $a < b < n$. We will first show that

$$n < a + b < n + 2 \frac{\log n}{\log 2} + 4. \quad (1)$$

To see this, define $\alpha(k)$ by $2^{\alpha(k)} |k|$ and $2^{\alpha(k)+1} \nmid k!$. It is an elementary result that

$$\begin{aligned} \alpha(k) &= \sum_{j=1}^{\lfloor \log k / \log 2 \rfloor} \lfloor 2^{-j} k \rfloor \geq \sum_{j=1}^{\lfloor \log k / \log 2 \rfloor} (2^{-j} k - 1) \\ &> k - 2 - \left\lfloor \frac{\log k}{\log 2} \right\rfloor. \end{aligned}$$

Since $\alpha(n) = \alpha(a) + \alpha(b)$ we then have

$$\begin{aligned} a + b &< \alpha(a) + \frac{\log a}{\log 2} + 2 + \alpha(b) + \frac{\log b}{\log 2} + 2 \\ &< \alpha(n) + 2 \frac{\log n}{\log 2} + 4 \\ &< n + 2 \frac{\log n}{\log 2} + 4. \end{aligned}$$

Proving $n < a + b$ is straightforward.

Next we show

$$a \log a > (n - b) \log n. \quad (2)$$

To see this, note that $n! = a!b!$ implies that $a! = \prod_{i=0}^{n-b-1} (n - i)$, so that

$$\begin{aligned} n^{-(n-b)} a! &= n^{-(n-b)} \prod_{i=0}^{n-b-1} (n - i) = \prod_{i=0}^{n-b-1} \left(1 - \frac{i}{n}\right) \\ &> \prod_{i=0}^{a-1} \left(1 - \frac{i}{n}\right) > \prod_{i=0}^{a-1} \left(1 - \frac{i}{a}\right) = a^{-a} a!. \end{aligned}$$

This proves (2). (Note that $n < a + b$ from (1) was used in the proof.

We now suppose that there are infinitely many examples of $n! = a!b!$ with $n - b \geq 5 \log \log n$, since otherwise we have $c = 5$ with no further effort. We may also assume that n is sufficiently large. We now proceed to derive a contradiction. From (2) we have

$$a \log a > 5(\log n) \cdot (\log \log n)$$

for sufficiently large n . Since the inequality $a \leq 4 \log n$ would imply $a \log a \leq 5(\log n) \cdot (\log \log n)$ for large n , we must have

$$a > 4 \log n \quad (3)$$

for n large.

From (1) and (2) we have

$$a \log a > \left(a - 2 \frac{\log n}{\log 2} - 4\right) \log n$$

or

$$a \log \left(\frac{n}{a}\right) < \left(2 \frac{\log n}{\log 2} + 4\right) \log n,$$

which implies

$$a \log \left(\frac{n}{a}\right) < 3(\log n)^2 \quad (4)$$

for large n .

However, since

$$\left(\frac{d}{da}\right)^2 a \log \frac{n}{a} = -\frac{1}{a} < 0$$

and since $4 \log n < a \leq n/2$ by (3), convexity implies

$$\begin{aligned} a \log \left(\frac{n}{a}\right) &\geq \min \left(4 \cdot \log n \cdot \log \left(\frac{n}{4 \log n}\right), \frac{n}{2} \log \left(\frac{n}{n/2}\right)\right) \\ &> 3(\log n)^2, \end{aligned}$$

contradicting (4) for large n . This completes the argument, showing that $n - b < 5 \log \log n$ for n sufficiently large.

Editorial comment. The proposer remarked that it would be nice to obtain a bound of the form $n - b = o(\log \log n)$.

No other solutions were received.

A Characterization of the Orthocenter

E 3466 [1991, 852]. *Proposed by William Fenton, Bellarmine College, Louisville, KY.*

Suppose $\triangle ABC$ is given. If X is a point not on any of the lines BC, CA, AB let the lines AX, BX, CX meet these lines respectively in points A', B', C' . It is known (Miquel's Theorem) that the circles $AB'C', A'BC', A'B'C$ intersect in a point Y . Prove that $X = Y$ if and only if X is the orthocenter of $\triangle ABC$.

Solution by Robin J. Chapman, University of Exeter, Exeter, U.K. Assume first that X is the orthocenter of $\triangle ABC$. Consider the circle with diameter XA . As the angles $XB'A$ and $XC'A$ are right angles, then B' and C' both lie on this circle. Hence X lies on the circle $AB'C'$. Similarly X lies on the circle $A'BC'$ and $A'B'C$. Hence $X = Y$ as required.

Conversely assume that $X = Y$. This is equivalent to the assertion that X lies on the circles $AB'C', A'BC'$ and $A'B'C$. The lines BC, CA and AB divide the plane into seven regions. We divide into cases according to which region X lies in.

Suppose first that X lies in the interior of $\triangle ABC$. Let $\theta = \angle XA'B$. As the quadrilateral $XA'BC'$ is cyclic then $\angle XC'B = \pi - \theta$ and so $\angle XC'A = \theta$. Similarly $\angle XB'C = \theta$.

Let A, B and C have position vectors \mathbf{a}, \mathbf{b} and \mathbf{c} respectively, with respect to the point X . Then

$$\begin{aligned}\mathbf{a} \cdot (\mathbf{b} - \mathbf{c}) &= |XA| |BC| \cos \theta, \\ \mathbf{b} \cdot (\mathbf{c} - \mathbf{a}) &= |XB| |CA| \cos \theta, \\ \mathbf{c} \cdot (\mathbf{a} - \mathbf{b}) &= |XC| |AB| \cos \theta.\end{aligned}$$

Adding we get

$$0 = (|XA| |BC| + |XB| |CA| + |XC| |AB|) \cos \theta.$$

Hence $\cos \theta = 0$ and so $\theta = \pi/2$ and X is the orthocenter of $\triangle ABC$.

Now suppose that X lies in one of the regions reached from the interior of $\triangle ABC$ by crossing one of the sides of the triangle. We may assume that this side is BC . This region is divided into four parts by the lines L and M , where L is the line through B parallel to AC and M is the line through C parallel to AB . Note that X cannot lie on L or M as then there would be no point B' or C' . Suppose now that X lies in the interior of the triangle whose side are L, M and BC . Then C lies between A and B' . It follows that X , which lies between C and C' , is contained in the interior of $\triangle AB'C'$. But then X cannot lie on the circle defined by A, B' and C' , a contradiction. Otherwise without loss of generality we may assume that X and C lie on the opposite side of L . Now A' lies between B and C and B lies between X and B' . Hence A' lies in the interior of $\triangle XB'C$, and so the points X, A', B' and C cannot lie on a circle, again giving a contradiction.

Finally assume that X lies in one of the three regions adjacent to just one vertex of $\triangle ABC$. We may assume that this vertex is A . Then A lies in the interior of $\triangle XBC$. Let $\phi = \angle XA'C$. Now, as the angles $\angle XA'C$ and $\angle XB'C$ subtend the same arc of the circle through X, A', B' and C , we have $\angle XB'C = \phi$. Also as the quadrilateral $XB'AC'$ is cyclic then $\angle XC'A = \pi - \phi$ and so $\angle AC'C = \phi$. Let \mathbf{u}, \mathbf{v} and \mathbf{w} be the position vectors of X, B and C respectively, with respect to A . Then

$$\begin{aligned}\mathbf{u} \cdot (\mathbf{w} - \mathbf{v}) &= |AX| |BC| \cos \phi, \\ \mathbf{v} \cdot (\mathbf{w} - \mathbf{u}) &= |AB| |CX| \cos \phi \\ \mathbf{w} \cdot (\mathbf{u} - \mathbf{v}) &= |AC| |XB| \cos \phi.\end{aligned}$$

Hence

$$0 = (|AB| |CX| + |AC| |XB| - |AX| |BC|) \cos \phi.$$

If X is not the orthocenter of $\triangle ABC$ then $\cos \phi \neq 0$ and $|AB| |CX| + |AC| |XB| = |AX| |BC|$. But a classical theorem (see §24.1 of D. Pedoe, *A Course of Geometry*, Cambridge, 1970) states that this is only possible if X , A , B and C are either collinear or concyclic. By hypothesis these points are not collinear, and as A lies in the interior of $\triangle XBC$ they cannot be concyclic either. This contradiction shows that X is the orthocenter of $\triangle ABC$.

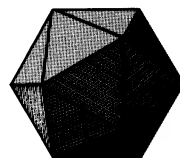
Editorial comment. Most of the submitted solutions are flawed by an incomplete (though easily completed) analysis of the case where X is outside $\triangle ABC$. A few solvers used coordinate methods to obtain straightforward, albeit messy, solutions. Jiro Fukuta noted that easy angle considerations show that if $X = Y$ then $\angle BXC$, is either congruent to or the supplement of $\angle BAC$, from which it follows that the circle through B , X , C is the reflection of the circumcircle of $\triangle ABC$ across side BC . Similarly, the reflections of the circumcircle across AB and CA also pass through X . The fact that X is the orthocenter then follows from the elegant “three circles theorem” of R. A. Johnson (see R. A. Johnson, *Modern Geometry*, Houghton Mifflin, 1929, p. 75), or the exposition on this and related theorems by D. N. Mackenzie, “Triquetras and Porisms,” *The College Math. J.*, 23 (1992), 118–131.

Solved also by E. Alkan (student, Turkey), J. Anglesio (France), F. Bellot & M. A. Lopéz (Spain), P. Čížek (student, Czech Republic), I. Dimitric, J. Fukuta (Japan), H. W. Guggenheimer, I. Kastanas, N. Komanda, J. H. Lee (Korea, student), H. Lipman, O. P. Lossers (The Netherlands), H. M. Marston, A. Nijenhuis, C. G. Petalas (Greece), V. Prasolov (Russia), P. S. Srinivasu (student, India), J. C. Vera (Colombia), M. Vowe (Switzerland), P. Zhao, Indian Institute of Technology Problem Group (India), and the proposer. Two incorrect and one incomplete solution were also received.

Collaborating editors: David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filasta, Ira M. Gessel, Richard A. Gibbs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pfiefer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, Edward T. H. Wang, and William E. Watkins.

ANSWER to picture puzzle: Bartel. L. van der Waerden
See p. 393.

The American Mathematical Monthly



Volume 100, Number 5 / MAY 1993



AN OFFICIAL PUBLICATION OF THE MATHEMATICAL ASSOCIATION OF AMERICA

NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

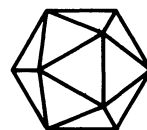
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International,
Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

The American Mathematical Monthly

Volume 100 Number 5 / MAY 1993
(ISSN 0002-9890)



Contents

ARTICLES

Thomas Archer Hirst—Mathematician Xtravagant I. A Yorkshire Surveyor
/ J. HELEN GARDNER and ROBIN J. WILSON 435

Hyperbolic Geometry on a Hyperboloid / WILLIAM F. REYNOLDS 442

A Simple Heuristic Proof of Hardy and Littlewood's Conjecture B /
MICHAEL RUBINSTEIN 456

The Pompeiu Problem / H. TURNER LAQUER 461

A Quicker Convergence to Euler's Constant /
DUANE W. DeTEMPLE 468

The Minimal Polynomial of $\cos(2\pi / n)$ / WILLIAM WATKINS and
JOEL ZEITLIN 471

The Equal Area Zones Property / B. RICHMOND and
T. RICHMOND 475

Graph Theory and the Game of Sprouts / MARK COPPER 478

FEATURES

COMMENTS 434

NOTES 483

PICTURE PUZZLE 488

THE AUTHORS 489

LETTERS 491

COMPUTER SCIENCE SAMPLER

Data Compression / CATHERINE C. McGEOCH 493

PROBLEMS AND SOLUTIONS 498

REVIEWS

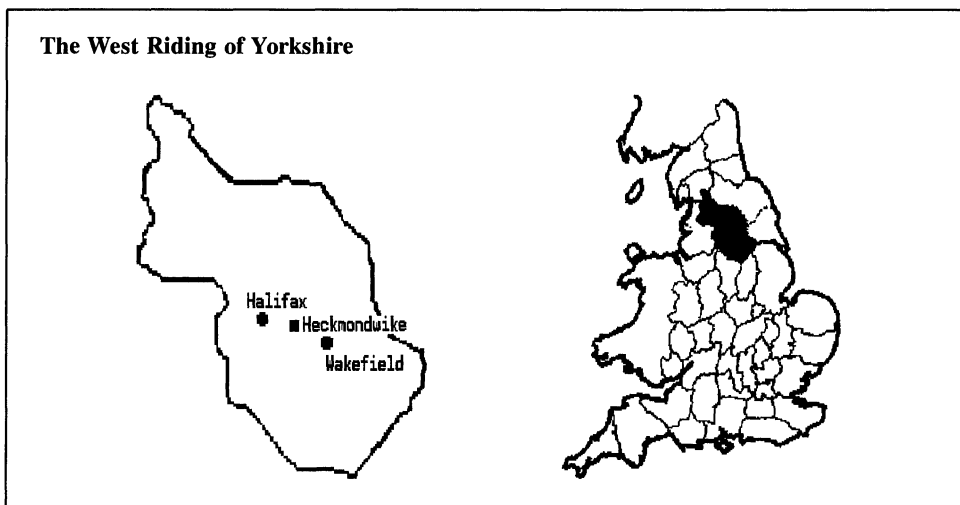
Linear Algebra Through Geometry by Thomas Banchoff and John Wermer /
JAMES KUZMANOVICH 506

TELEGRAPHIC REVIEWS 509

Thomas Archer Hirst— Mathematician Xtravagant I. A Yorkshire Surveyor

J. Helen Gardner and Robin J. Wilson

I was born at a small town in the West Riding of Yorkshire (England) called Heckmondwike . . . I was the youngest of four brothers . . . and my father having been successful in business and wishful to give his sons the advantage of the best education which the neighbourhood could afford, he removed to the vicinity of Wakefield. I was at this time but five years of age, and up to my eleventh year was sent to a preparatory school in Wakefield, after which, however, I became enrolled as a pupil of the West Riding Proprietary School, one of the first of a number of improved schools which at this period were founded in England . . . At this school, or, indeed, at any I remained but four years longer, during which period I could of course obtain the most rudimentary and necessary instruction. I remember, however, that here Mathematics was my favourite study and a promising opportunity of applying them now occurring, I was somewhat suddenly and, as I now think, prematurely, taken away from school and articled for five years to an Engineering Surveyor in Halifax about fifteen miles distant from Wakefield. This occurred in August 1845, a period celebrated in English history as the commencement of the "Railway Mania." . . .



In spite of this commonplace start, Thomas Archer Hirst pursued a most distinguished career. He was awarded a Ph.D. in geometry in Marburg, Germany, in 1852; he met Gauss and Weber in Göttingen and studied in Berlin with Dirichlet and Steiner; he became acquainted with Poincaré, Liouville, and Chasles in Paris, and with Brioscchi, Tortolini, and Cremona in Italy; back in England, he

The Diaries of Thomas Hirst

Thomas Hirst's candid observations on his contemporaries reveal the personalities behind the names:

Arthur Cayley

...a thin weak-looking individual with a large head and face marked with small-pox; he speaks with difficulty and stutters slightly.



Carl Friedrich Gauss

...a venerable, fine old fellow, with a contented manly expression. He is about 80 years of age, but not a trace of superannuation is to be seen about him.

Michael Faraday

...such soundness, such cheering freshness and want of pretension—I sat in silent admiration which increased with every moment. He is a fine fellow, a beautiful character.



Augustus De Morgan

A dry dogmatic pedant I fear is Mr de Morgan, notwithstanding his unquestioned ability.

was a schoolmaster in Hampshire and London, and became a pioneer in the teaching of geometry; he became a Fellow of the Royal Society at the age of 31, served on its Council for many years, and was awarded the Royal Society medal; he was a close associate of Cayley, Sylvester and De Morgan, and also knew Stokes, Maxwell, Boole, Faraday and Darwin; he was a founder member of the X-club, a small group of distinguished scientists including Huxley and Tyndall, from which the nickname in the title derives; he was the first Vice-President, and later President, of the London Mathematical Society; he was Professor of Mathematical Physics, and later (following De Morgan) Professor of Mathematics at University College, London; he became General Secretary of the British Association and Assistant Registrar of the University of London; he was a co-founder, and the first President of, the Association for the Improvement of Geometrical Teaching, later to become the Mathematical Association; and he became Director of Studies at the Royal Naval College at Greenwich, where he entertained Chebyshev and Klein. He died in London on 16 February 1892 at the age of 62.

Yet despite all these achievements, Thomas Hirst would have become a forgotten figure had it not been for his habit of writing a journal; and just as the diaries of Samuel Pepys and John Evelyn describe the London of the 1660s, so we can learn much about scientific life in the Victorian era from the extensive diaries he kept. Covering over forty-five years, from 1845, when he moved to Halifax as an apprentice of 15, to 1892, a month before his death, and extending to almost 3000 pages of typescript, they chronicle with great clarity, if pomposity at times, the scientific circles in which he moved—both in England and Europe.

Thomas Archer Hirst was born into a middle-class home on 22 April 1830. His parents were both from successful wool-merchant families—indeed, when Hirst's maternal grandfather died, his father, a wool-stapler, was wealthy enough to be able to retire at the early age of 31. The death of Thomas's father from a drinking accident in 1844 thrust on his mother the responsibility of finding careers for her sons, and she arranged for Thomas, then aged 15, to be articled to the engineer Richard Carter, who had been commissioned to survey from Halifax to Keighley for the West Yorks Railway.

It was with this move to Halifax that Hirst began the record of his adult life:

25th August 1845: Came to Halifax. Got here at 12.30 o'clock. Afternoon at office writing. Evening with John at Mr Richardson's, playing music. Mr Carter from home. Wrote to Mother...

and he quickly settled in:

31st August 1845: Sunday. Morning at Chapel with John. Afternoon at Mr Tyndall's to tea—present Messrs Tyndall, Richardson, Tidmarsh, brave John and I. Evening at old Church. Sent letter to Will. Walking, singing with Miss Carter for first time.

The 'John' was his brother and the 'Mr Tyndall' was a young Irishman, John Tyndall, Carter's principal surveyor and ten years Hirst's senior. This was the beginning of a close friendship that was to have a great influence on Hirst's life. As he later wrote:

15th August 1848: There is no person of whose acquaintance I feel more proud, no friend for whom I have more regard, and for whose abilities I have more esteem. May that acquaintance, from which I have already reaped so much, and from which I hope to reap more, continue...

19th-century Halifax
(a contemporary print by J. R. Smith.)



The admiration seems to have been mutual, for Tyndall describes Hirst in his own diary as

“our junior apprentice, a youth upwards of 6 feet high and about 16 years of age—an immense development of brain which is in true keeping with his extraordinary powers of thinking.”

And Tyndall was not the only notable Victorian who was charmed by him. Three years later, the radical essayist and poet January Searle, wrote:

“He is a fine fellow ... a man measuring six feet two inches of perpendicular flesh, of excellent proportions, and a face much handsomer than any I see here. ... Just such is my friend Tom. ... He is a generous warm-hearted fellow, but for the most part silent and reserved in company, especially if there be anything to learn. He is a musician likewise, and when I go to see him we always have a happy time, and discuss subjects of serious import, which, I am sorry to say, are not popular in mixed company. You would love Tom, for he has a most affectionate nature, although he does not show it in words”

But he was not always silent and reserved in company. Even at the age of 17 he enjoyed his pipe and cigars, and a glass of porter or whiskey.

30th April 1847: Playing, singing, etc. etc. all the evening ... until about 5 o'clock, then took the whiskey bottles to bed. Bill got dead drunk, and then began spewing on to my best breeches. Have a confused recollection of having an argument with Mr Garces on the influence of wine, but how it ended God knows.

His work kept him very busy, and he worked long hours—both in the office, drawing and tracing plans, out in the field, levelling and surveying, and constantly travelling from place to place. The scenery delighted him and his diary is littered with descriptions of places that had left impressions in his mind. Away from work, he found much to interest him in the press, often commenting on world events such as the famine in Ireland, the 1848 revolution in France, and the discovery of gold in California.

Inspired by Tyndall, Hirst started on a programme of self-improvement. He read widely—all the latest novels, such as *Oliver Twist* and *Jane Eyre*—and extensively on science topics. In mathematics, he read geometry from Euclid's *Elements*, trigonometry and algebra from Hutton's *Mathematics*, and Sir David Brewster's *Life of Sir Isaac Newton*. Around this time, the diaries become more discursive. Anything which affected him was recorded in fine detail, with long accounts of lectures, sermons, discussions, books, and events. Gradually, he began to show an uncompromising attitude to study and reading:

16 July 1847: Did not get up until 8.30, so therefore missed my French lesson, which showed me the necessity of not giving too much to pleasure. I don't think I felt quite so comfortable this morning as I should have done had I devoted yesterday evening to Lyell's geology—thus showing that study is the only pleasure unattended in some degree with remorse.

There are entries on state education, on political issues such as capital punishment and the issue of Church and State, on morals, philosophy and literature, and on the relationship between science and religion:

1st September 1847: Reading Joyce on "The Seasons", ... and trying to picture in my mind the earth, revolving on its axis, the moon revolving round the earth, and the earth together with the moon, revolving at the same time is enough to make one's wits themselves revolve. Everything you set your eyes on seems to be a moon revolving. You fancy the chairs, table, and everything in the room revolving round the candle as a sun—until at length you are driven to close the book in utter despair. Such was the situation I found myself in this evening; the more I read, the more mysteries rose up before me. What an awful example of our insignificance.

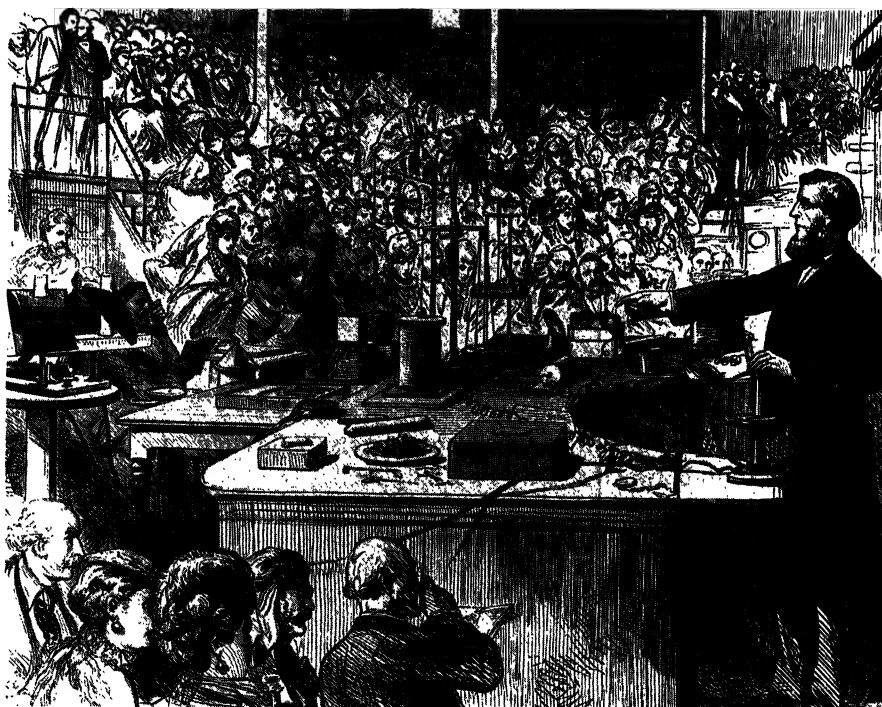
This was a time when general science and education were beginning to establish themselves as activities for the general populace. In 1824 the first series of lectures was given at the London Mechanics Institute, and by 1850 there were 610 Mechanics Institutes throughout England, as well as numerous Improvement Societies. Hirst first enrolled at the Halifax Mechanics Institute in February 1848, and became increasingly involved in it; but although he "taught a little at the Mathematical Class", he was having difficulties in deciding the direction of his own studies:

1st June 1848: Reading Hutton's *Mathematics*, *Arithmetical and Geometrical Progression* and *Surds*, I begin to find the want of a regular system of mathematical reading; I have done a little here and there before, but mastered none.

Meanwhile, John Tyndall had left Halifax, undertaking a short spell of teaching at Queenwood College in Hampshire, before going to study chemistry with his friend and colleague Edward Frankland at the University of Marburg in Germany. Among the small but distinguished faculty at Marburg was Robert Bunsen, inventor of the bunsen burner. Hirst went out there in August 1849 to visit him, and his diaries record how he enjoyed the stimulus of Tyndall's conversation, as well as meeting Tyndall's friends, sitting in on lectures, and soaking up the scenery.

John Tyndall (1820–1893)—‘a queer, independent, and perhaps also partly insane man’

Tyndall was one of the most celebrated popularizers of Victorian science. In 1853 he became professor of natural philosophy at the Royal Institution, working with, and eventually succeeding, Faraday. His celebrated ‘Belfast address’ on science and religion, given to the British Association in 1874, brought him into direct conflict with the Church. His research ranged from magnetism and heat radiation to glacial movement and bacteriology, and he was the first to explain why the sky is blue.



Regrettably, the visit had suddenly to be cut short by the unexpected death of his mother at the age of 44. The turmoil of his mother's death led Hirst to assess his position, not only at home but at work, since he was now financially more independent. On Tyndall's departure he had been promoted to Carter's chief surveying assistant, but now he became increasingly convinced that surveying was not the career he sought. As he wrote three years later:

This choice, then apparently so promising, proved otherwise. The first three years of my apprenticeship, it is true, were passed in great activity, but not fifty per cent of the railway schemes then projected were ever constructed; immense sums of money were expended and the profession so overcrowded that when the reaction came, many—among whom were Doctor Tyndall and myself—found it expedient to leave the profession.

Having considered all things he looked towards a different future:

30th September 1849: ...I have carved out for myself a path, a duty: I have contemplated calmly an *ideal future*. Heaven strengthen me in my resolve to make it an *actual present*. I will

not particularize it—perhaps I could not; but the idea will remain there as my guiding star... Away now for ever with the dream of outward support and guidance. I alone have grasped the helm that is to guide me through this world, and if at last I reach the haven, to me alone will belong the honour.

The details of this passage remain a mystery, but four months later he wrote:

26th January 1850: Reading algebra. My close attention to mathematics for two or three days has made me long for continued opportunities for its study, and almost determined me so to endeavour.

Despite this resolve, Hirst prudently completed his apprenticeship. On 31st August 1850, he said goodbye to surveying for ever. Attracted by Tyndall's experience in Germany, he returned to Marburg to study mathematics, physics and chemistry. His time as a student in Marburg will be the topic of the next article.

ACKNOWLEDGMENTS. A typed version of the Thomas Hirst diaries is held at the Royal Institution in London, and quotations from the diaries appear here by courtesy of the Royal Institution. The diaries have been edited by W. H. Brock and R. M. MacLeod, and were published in microfiche by Mansell, London, in 1980. We are grateful to Professor Brock for his help on many occasions.

Open University
Milton Keynes MK7 6AA
England

On January 19, 1913, occurred in Denver the death of Robert Gauss, who had been long connected with the *Denver Republican*. On the same day died also his brother, Charles H. Gauss, of St. Louis. They were sons of Eugene Gauss, and grandsons, of the mathematician, CARL FRIEDRICH GAUSS.

—*American Mathematical Monthly*
20, (1913) p. 71

Hyperbolic Geometry on a Hyperboloid

William F. Reynolds

1. INTRODUCTION. Hardly anyone would maintain that it is better to begin to learn geography from flat maps than from a globe. But almost all introductions to hyperbolic non-Euclidean geometry, except [6], present plane models, such as the projective and conformal disk models, without even mentioning that there exists a model that has the same relation to plane models that a globe has to flat maps. This model, which is on one sheet of a hyperboloid of two sheets in Minkowski 3-space and which I shall call H^2 , is over a hundred years old; Killing and Poincaré both described it in the 1880's (see Section 14). It is used by differential geometers [29, p. 4] and physicists (see [21, pp. 724–725] and [23, p. 113]). Nevertheless it is not nearly so well known as it should be, probably because, like a globe, it requires three dimensions.

The main advantages of this model are its naturalness and its symmetry. Being embeddable (distance function and all) in flat space-time, it is close to our picture of physical reality, and all its points are treated alike in this embedding. Once the strangeness of the Minkowski metric is accepted, it has the familiar geometry of a sphere in Euclidean 3-space E^3 as a guide to definitions and arguments. For example, the lines of H^2 are its non-empty intersections with the planes through the origin of the Minkowski space M^3 . The length of a line segment of H^2 is defined by analogy with arc length in calculus; this leads naturally to the hyperbolic functions. As in the spherical case, every isometry of H^2 can be extended to a linear transformation of M^3 , so that straightforward calculations with matrices can be used to prove theorems and develop the trigonometry of H^2 . The circles, horocycles, and equidistant curves have a beautiful interpretation: they are precisely the nontrivial intersections of H^2 with planes of M^3 that do *not* pass through the origin. An area function for H^2 can be constructed from the volume function on M^3 .

The aim of this article is to show that H^2 can be used to give an introduction to hyperbolic geometry to undergraduates who know a little about linear transformations and groups, a bit of special relativity being helpful for motivation. The mixing of rigor and intuition is similar to what is common in calculus courses. The treatment is not axiomatic, since there is no intrinsic reason to stress axiomatics in hyperbolic geometry any more than in, say, spherical trigonometry. For historical and philosophical reasons, however, many treatments are based on axioms; therefore I shall refer to Moise's axioms [22] at the points where they can be verified from my approach. (I have chosen these axioms since, incorporating real-valued functions for distance and measure of angles, they are closer to my analytic approach than Hilbert's [6], [8]; the latter, being weaker, would be easier to verify.) I will treat them not as axioms, but just as properties of H^2 . I have included something about the hyperbolic analogues of map projections and about the history of the model.

A special feature of my approach, which distinguishes it from Faber's [6, Chapter VII], is its extensive use of orthogonal transformations of M^3 , the analogues of the rigid motions of E^3 that fix the origin, to move subsets of H^2 to convenient positions.

I want to thank Alan H. Durfee and Mark E. Kidwell for encouraging me to write up this article. It originated in an undergraduate course at Tufts, and I did some of the work while visiting at Harvard.

2. MINKOWSKI 3-SPACE. By *Minkowski 3-space* M^3 I mean a 3-dimensional real vector space together with a real-valued function q on it such that

$$q\left(\sum_{i=0}^2 x_i U_i\right) = -x_0^2 + x_1^2 + x_2^2 \quad (2.1)$$

for some basis $\mathcal{U} = (U_0, U_1, U_2)$ of the space. More briefly,

$$q\left(\sum x_i U_i\right) = \sum e_i x_i^2 \quad (2.2)$$

where $e_0 = -1$, $e_1 = e_2 = 1$; then $q(U_i) = e_i$. We can define *Minkowski n -space* M^n similarly for any $n \geq 2$ (with one minus sign). To relate the cases $n = 2$ and 3, take M^2 to be the subspace of M^3 with basis (U_0, U_1) together with the restriction of q .

If we replace (2.2) by

$$q_E\left(\sum x_i U_i\right) = \sum x_i^2, \quad (2.3)$$

we get ordinary Euclidean 3-space E^3 , with q_E giving the square of length. I shall constantly use analogies from E^3 to study M^3 ; watch for such analogies when they are not mentioned.

For $X = \sum x_i U_i$ and $Y = \sum y_i U_i$ in M^3 , define

$$p(X, Y) = \frac{1}{2}[q(X + Y) - q(X) - q(Y)]. \quad (2.4)$$

(This is the *bilinear form* or *pairing* corresponding to the *quadratic form* q .) Then

$$q(X) = p(X, X) \quad (2.5)$$

and

$$p(X, Y) = \sum e_i x_i y_i = -x_0 y_0 + x_1 y_1 + x_2 y_2; \quad (2.6)$$

in particular $p(U_i, U_j)$ equals e_i if $i = j$ and 0 otherwise. Observe that p is analogous to the ordinary dot product on E^3 given by $p_E(X, Y) = X \cdot Y = \sum x_i y_i$ and \mathcal{U} to an orthonormal basis.

3. THE HYPERBOLOIDAL MODEL. In M^3 , let H^2 be the set of all vectors $X = \sum x_i U_i$ for which

$$q(X) = -1, \quad (3.1)$$

$$x_0 > 0. \quad (3.2)$$

These two conditions can be expressed by the equation

$$x_0 = \sqrt{1 + x_1^2 + x_2^2}. \quad (3.3)$$

(3.1) describes a hyperboloid of two sheets and (3.2) picks out one sheet. We can define $H^n \subset M^{n+1}$ similarly for arbitrary n ; for $n = 1$, take $H^1 = H^2 \cap M^2$ with M^2 as in the previous section. FIGURE 1 may help in thinking about H^2 . In flat

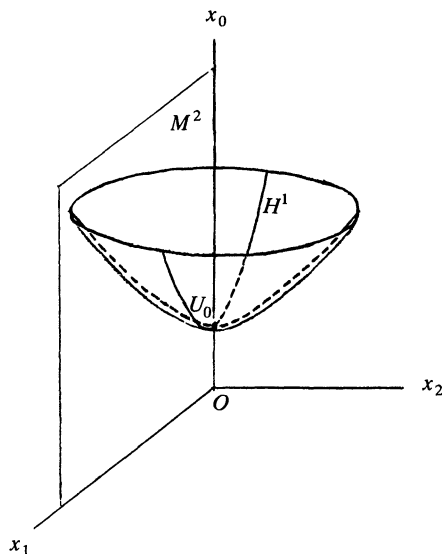


Figure 1

space-time, identified with M^4 , H^2 appears to each observer as a circle whose radius is increasing at slightly greater (!) than the speed of light.

We now begin to construct a model of hyperbolic geometry whose *points*, or *H-points*, are the elements of H^2 . We think of them as either points or vectors when considered in M^3 , and as points when considered in H^2 . The *lines* or *H-lines* of the model are defined to be all the nonempty intersections of H^2 with 2-dimensional subspaces of M^3 ; for example, H^1 is an *H-line*. (The prefix “H-” will always be optional.)

Incidence of *H-points* and *H-lines* and *betweenness* for points of an *H-line* are defined in the natural way. Each pair of distinct *H-points* A and B lies on a unique *H-line* \overleftrightarrow{AB} , namely the intersection of H^2 with the plane OAB of M^3 , O being the origin of M^3 . The definitions of the *H-segment* \overline{AB} and the *H-ray* \overrightarrow{AB} are straightforward. We can now check the plane incidence axioms of [22, pp. 37–38].

It should be clear that the complement of each *H-line* in H^2 consists of two *H-half planes*, called its *sides*. This statement can be made precise as the plane-separation axiom or Pasch’s axiom [22, p. 62]. Two distinct *H-lines* have one or zero *H-points* in common according as the line of intersection of the planes of M^3 in which they lie intersects H^2 or not; this gives the hyperbolic parallel axiom [22, p. 114]. The Archimedean axiom and the axiom of completeness (or continuity) [22, pp. 256 and 265] are also clear.

In the analogous situation in E^3 , the equation $q_E(X) = 1$ (cf. (2.3)) defines a sphere S^2 ; this leads to a model of spherical (or double elliptic) geometry whose lines or *S-lines* are the great circles of S^2 .

4. DISTANCE. For distinct *H-points* A and B we want to define the *H-distance* $d(A, B)$, also called the *H-length* of \overline{AB} . The natural way to adapt the usual definition of arc length is to partition \overline{AB} , as a curve in M^3 , by suitable points

$P_0 = A, P_1, \dots, P_m = B$ and to define

$$d(A, B) = \lim_{m \rightarrow \infty} \sum_{j=1}^m \sqrt{q(P_j - P_{j-1})}; \quad (4.1)$$

but first we must check that $q(P_j - P_{j-1})$, analogous to a squared length, is positive.

The equation $q(X) = 0$ describes a cone in M^3 and the vectors with $q(X) > 0$ are the points outside this cone (cf. the spacelike vectors of special relativity). By the mean value theorem there is a point of $\overline{P_{j-1}P_j}$ at which the tangent vectors in M^3 to this H -segment are parallel to $P_j - P_{j-1}$, so it suffices to show that all vectors $V \neq O$ tangent to H^2 have $q(V) > 0$. We can show this by turning to advantage a limitation of our intuition. (For a different approach, see Section 7.) Since we are used to Euclidean space, any attempt to visualize M^3 pictorially as in FIGURE 1 identifies it with E^3 , i.e. imposes a Euclidean quadratic form on it. This destroys the symmetry of M^3 and of H^2 , so that different points of H^2 do not look alike. Since we cannot avoid this identification, we use it. Suppose then that $M^3 = E^3$ has both forms q and q_E with respect to the basis \mathcal{Q} . The cone $q(X) = 0$ now consists of all vectors that make (Euclidean) angles of $\pi/4$ with the plane $x_0 = 0$. All tangent vectors V to H^2 in E^3 make angles less than $\pi/4$ with that plane, so that $q(V) > 0$; this implies that the definition (4.1) of $d(A, B)$ is valid as in the Euclidean case.

To evaluate the limit, parametrize \overline{AB} ; that is, let F be a smooth one-to-one mapping of an interval $a \leq t \leq b$ onto \overline{AB} with $F(a) = A, F(b) = B$. F is a vector-valued function of t with derivative F' and $P_j = F(t_j)$ for a partition $a = t_0 < t_1 < \dots < t_m = b$. The limit is

$$d(A, B) = \int_a^b \sqrt{q(F'(t))} dt. \quad (4.2)$$

Moise's axioms of distance and definition of betweenness [22, pp. 47–49 and 51] can be obtained now; also see Section 12. For *curves* in H^2 , (4.1) and (4.2) give *H-arc length*.

We can use (4.2) to compute H -distances along H^1 ; we shall treat H^2 in Section 6. Let \overline{AB} be any H -line segment of H^1 , say $A = a_0U_0 + a_1U_1, B = b_0U_0 + b_1U_1$, with $a_1 < b_1$. Using (3.3) we can take the parameter $t = x_1$, with $F(t) = \sqrt{1+t^2}U_0 + tU_1, a_1 \leq t \leq b_1$. Then

$$d(A, B) = \int_{a_1}^{b_1} \sqrt{-\frac{t^2}{1+t^2} + 1} dt = \left[\ln(t + \sqrt{t^2 + 1}) \right]_{a_1}^{b_1}.$$

Define

$$\operatorname{arcsinh} t = \ln(t + \sqrt{t^2 + 1});$$

$\operatorname{arcsinh}$ is monotone increasing and $\operatorname{arcsinh} 0 = 0$. Then

$$d(A, B) = \operatorname{arcsinh} b_1 - \operatorname{arcsinh} a_1;$$

in particular $d(U_0, B) = \operatorname{arcsinh} b_1$ if $b_1 > 0$, whence the H -length of H^1 is infinite. If $r = \operatorname{arcsinh} t$, then $t = (e^r - e^{-r})/2 = \sinh r$, the *hyperbolic sine* of r . On H^1 , $x_0 = \sqrt{1+t^2} = (e^r + e^{-r})/2 = \cosh r$, the *hyperbolic cosine* of r ; so a parametrization of H^1 by H -length is

$$P(r) = (\cosh r)U_0 + (\sinh r)U_1, \quad -\infty < r < \infty. \quad (4.3)$$

I use the slightly unusual symbol “ $\operatorname{arcsinh} t$ ” since it does represent an (arc) length in our peculiar way of measuring.

5. ORTHOGONAL TRANSFORMATIONS. A linear transformation T of M^3 to M^3 is called *orthogonal* (with respect to q) if

$$q(T(X)) = q(X), \quad X \in M^3. \quad (5.1)$$

The orthogonal transformations form a group, the *orthogonal group* $\mathbf{O}(M^3)$.

We can express this in terms of matrices as follows. Each vector $X = \sum x_i U_i$ of M^3 has its column of coordinates

$$[X] = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \quad (5.2)$$

with respect to the basis \mathcal{U} . Each linear transformation T has a matrix $[T] = [t_{ij}]$ with respect to \mathcal{U} such that the matrix equation

$$[T(X)] = [T][X] \quad (5.3)$$

holds for all X . By (2.4) and (2.5), T is in $\mathbf{O}(M^3)$ if and only if

$$p(T(X), T(Y)) = p(X, Y), \quad X, Y \in M^3, \quad (5.4)$$

or, equivalently,

$$p(T(U_j), T(U_k)) = \sum_{i=0}^2 e_i t_{ij} t_{ik} = \begin{cases} e_j & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \quad (5.5)$$

(since $[T(U_j)]$ is the j -th column of $[T]$). (5.5) can be written as

$$[T]^t [J_0] [T] = [J_0], \quad (5.6)$$

where $[T]^t$ is the transpose of $[T]$ and

$$[J_0] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

By (5.6), the determinant of T is ± 1 . The elements of $\mathbf{O}(M^3)$ with determinant 1 form a subgroup $\mathbf{O}^+(M^3)$ of index 2, the *special orthogonal group*.

The orthogonal group has a subgroup associated to H^2 that will be fundamental to us. Let $T \in \mathbf{O}(M^3)$; if $X \in H^2$, (3.1) and (5.1) imply that $T(X) \in H^2$ or $-T(X) \in H^2$. By continuity, those T that map H^2 onto itself form another subgroup of $\mathbf{O}(M^3)$ of index 2, which I shall call $\mathbf{G}(M^3)$ or simply \mathbf{G} . Furthermore $\mathbf{G}(M^3)$ has the subgroup $\mathbf{G}^+(M^3) = \mathbf{G}(M^3) \cap \mathbf{O}^+(M^3)$ of index 2.

Let $T \in \mathbf{G}(M^3)$; since T maps each 2-dimensional subspace of M^3 onto another, it maps each H -line onto another. Then, by (4.1) and (5.1), if A and B are points of H^2 ,

$$\begin{aligned} d(T(A), T(B)) &= \lim \sum \sqrt{q(T(P_j) - T(P_{j-1}))} \\ &= \lim \sum \sqrt{q(T(P_j - P_{j-1}))} = \lim \sum \sqrt{q(P_j - P_{j-1})} = d(A, B). \end{aligned} \quad (5.7)$$

This means that the restriction to H^2 of every element of $\mathbf{G}(M^3)$ is an H -isometry or distance-preserving mapping of H^2 .

Analogous groups exist for M^n ; for example, $\mathbf{G}^+(M^4)$ is the Lorentz group of special relativity. Similarly, in Euclidean space, $\mathbf{O}(E^3)$ consists of the usual orthogonal transformations and $\mathbf{O}^+(E^3)$ of the rotations about the origin. (Since there is no analogue of the sheets of H^2 , $\mathbf{O}(E^3)$ corresponds to $\mathbf{G}(M^3)$ as well as to $\mathbf{O}(M^3)$.)

6. THE FORMULA FOR DISTANCE. The following calculations, first over M^2 and E^2 and then in $\mathbf{G} = \mathbf{G}(M^3)$, will motivate equations (6.4) and (6.7); some readers may prefer to skip them or to use the alternative argument indicated at the end of this section. By the two-dimensional analogue of (5.5), it is a simple exercise to show that $T \in \mathbf{O}(M^2)$ if and only if

$$[T] = \begin{bmatrix} e \cosh s & f \sinh s \\ e \sinh s & f \cosh s \end{bmatrix}, \quad e = \pm 1, \quad f = \pm 1, \quad (6.1)$$

with e, f, s uniquely determined by T . Then $T \in \mathbf{G}(M^2)$ if and only if $e = 1$, $T \in \mathbf{O}^+(M^2)$ if and only if $e = f$, and $T \in \mathbf{G}^+(M^2)$ if and only if $e = f = 1$. Similarly, as is well known, $\mathbf{O}(E^2)$ consists of those T such that

$$[T] = \begin{bmatrix} \cos \theta & -h \sin \theta \\ \sin \theta & h \cos \theta \end{bmatrix}, \quad h = \pm 1. \quad (6.2)$$

$T \in \mathbf{O}^+(E^2)$ if and only if $h = 1$.

Now let \mathbf{G}_1 be the subgroup of \mathbf{G} consisting of those T that fix the H -line H^1 (as a whole), i.e. such that $T(H^1) = H^1$ in the usual notation. If $T \in \mathbf{G}_1$, then T also fixes M^2 , so that $t_{20} = t_{21} = 0$. By (5.5), it is easy to see that $t_{02} = t_{12} = 0$, i.e., T fixes the subspace spanned by U_2 . (This is a special case of a fact about orthogonal complements [11, p. 364].) On M^2 , T acts like an element of $\mathbf{G}(M^2)$; then, by (6.1),

$$T = L_s J_1^i J_2^j, \quad i, j = 0, 1, \quad (6.3)$$

where

$$[L_s] = \begin{bmatrix} \cosh s & \sinh s & 0 \\ \sinh s & \cosh s & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.4)$$

for some real s ,

$$[J_1] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad [J_2] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Thus \mathbf{G}_1 is the set of all transformations of the form (6.3). $L_s L_t = L_{s+t}$ and $L_0 = I$, the identity transformation.

Similarly, if T is in the subgroup \mathbf{G}_0 of elements of \mathbf{G} that fix U_0 , T fixes the subspace spanned by U_1 and U_2 . Since

$$q(x_1 U_1 + x_2 U_2) = x_1^2 + x_2^2, \quad (6.5)$$

this subspace with the restriction of q is a Euclidean plane. By (6.2),

$$T = R_\theta J_2^j, \quad j = 0, 1, \quad (6.6)$$

where

$$[R_\theta] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}; \quad (6.7)$$

thus G_0 consists of all T of the form (6.6). Note that in E^3 , identified with M^3 as in Section 4, R_θ is a rotation through θ radians. Clearly

$$G_1 \cap G_0 = \{I, R_\pi, J_1, J_2\}. \quad (6.8)$$

Applying R_θ to the parametrization (4.3) of H^1 , define $P(r, \theta) = R_\theta(P(r))$. By (5.3),

$$P(r, \theta) = (\cosh r)U_0 + (\sinh r \cos \theta)U_1 + (\sinh r \sin \theta)U_2. \quad (6.9)$$

This is a parametrization of H^2 . It is true that r and θ are *hyperbolic polar coordinates* for H^2 ; the meaning of r follows from (4.3) since R_θ is an isometry, but that of θ must wait until we define H -angle measure in the next section. By matrix multiplication, $L_s(P(r, 0)) = P(r + s, 0)$ and $R_\theta(P(r, \phi)) = P(r, \phi + \theta)$. Thus we now call L_s the H -translation by s along H^1 and we will call R_θ the H -rotation by θ about U_0 .

At last we can obtain the basic formula for H -distance, namely:

$$d(A, B) = \operatorname{arccosh}(-p(A, B)). \quad (6.10)$$

To prove this, note that by (5.7) and (5.4) both sides are unchanged by applying any element of G . Applying $L_{-r}R_{-\theta}$ where $A = P(r, \theta)$, we can reduce to the case $A = U_0$; applying another R_ϕ , we can further assume that $B = P(s, 0)$ with $s > 0$. By (2.6) and (4.3), $p(U_0, B) = -\cosh s$; since $s = d(U_0, B)$, this implies (6.10).

(6.10) may be compared with the fact that in spherical geometry the distance $d_S(A, B)$ between two points of S^2 is the radian measure of the Euclidean angle $\angle AOB$, which is $\arccos p_E(A, B)$.

Most of the above argument was devoted to the reduction to the case $A = U_0$, i.e. the proof of the basic fact that *every point A of H^2 can be moved to U_0 by an element of G* ; this means that the points of H^2 are "all alike" as far as G is concerned. At the cost of using more linear algebra, we could have eliminated some calculations by proving this as follows. By a slight variant of the Gram-Schmidt orthogonalization process [11, pp. 356–357] together with Sylvester's Theorem [11, p. 359], we could show that the coordinate column $[A]$ of A (cf. (5.2)) is the first column of the matrix of some element T of G ; then T^{-1} is the required element.

7. ANGLES. The H -angle $\angle BAC$ is $\overrightarrow{AB} \cup \overrightarrow{AC}$ where $\overrightarrow{AB} \neq \overrightarrow{AC}$. Define its (H -angle) *measure* $m(\angle BAC)$ as follows: let V and W be the vectors in M^3 tangent to \overrightarrow{AB} and \overrightarrow{AC} respectively at A such that $q(V) = q(W) = 1$; then set

$$m(\angle BAC) = \arccos p(V, W). \quad (7.1)$$

To see that these vectors exist and are unique and that $|p(V, W)| \leq 1$, reduce to the case $A = U_0$ by transforming everything by a suitable element of G as above; then the tangent plane to H^2 is $x_0 = 1$, which is Euclidean by (6.5), so that (7.1) makes sense. In fact, the ordinary law of cosines shows that the H -angle measure of $\angle BU_0C$ is the same as its Euclidean measure, so that in (6.9) θ really measures an H -angle. The angle measure between two intersecting curves in H^2 is defined similarly. Clearly every element of G preserves angle measure just as it preserves distance. (A similar argument could have been used in Section 4 to show that $q(V)$ is positive.)

8. SUPERPOSITION. Now we have all we need to prove the following main existence and uniqueness theorem for elements of G . This theorem tells us exactly how far we can extend the fact that every point of H^2 can be moved to U_0 by some element of G .

Theorem. Let l_j be an H -line, $\overrightarrow{A_j B_j}$ a ray on l_j , and S_j a side of l_j , for $j = 1, 2$. There exists exactly one $T \in \mathbf{G}$ such that $T(A_1) = A_2$, $T(l_1) = l_2$, $T(\overrightarrow{A_1 B_1}) = \overrightarrow{A_2 B_2}$, and $T(S_1) = S_2$.

To see this, note that since \mathbf{G} is a group, we can assume, as in the proof of (6.10), that $\overrightarrow{A_j B_j}$ is the ray of H^1 for which $x_1 \geq 0$ for $j = 1, 2$. Then by (6.8) T is J_2 or I , depending on whether or not it interchanges the sides of H^1 ; the theorem follows.

Let us call two subsets K and K' of H^2 *congruent* if $T(K) = K'$ for some $T \in \mathbf{G}$. Then the theorem implies that two H -segments are congruent if and only if they have the same H -length, and that two H -angles are congruent if and only if they have the same measure. The axioms for angle-construction, angle-addition, and supplements [22, pp. 76–77] and the SAS-axiom [22, p. 84] can now be verified easily. This completes Moise's list of axioms and shows that H^2 is indeed a model for plane hyperbolic geometry.

We could now establish that every isometry of H^2 is the restriction of some element of \mathbf{G} , using the theorem together with the SSS-theorem [22, p. 87]. (Cf. the corresponding fact for S^2 and $\mathbf{O}(E^3)$.) Thus two subsets of the hyperbolic plane are congruent if and only if one can be mapped on the other by “superposition”, i.e. by applying an isometry. This condition does not depend on the particular model of the hyperbolic plane being used.

9. TRIGONOMETRY. Let $\triangle ABC$ be any triangle in H^2 ; as usual in trigonometry, set $d(B, C) = a$, $d(C, A) = b$, $m(\angle BAC) = \alpha$, etc. Transforming by a suitable element of \mathbf{G} we can suppose that $C = U_0$, that A is on the ray of H^1 for which $x_1 \geq 0$, and that B is on the side of H^1 for which $x_2 > 0$; then $A = P(b, 0)$ and $B = P(a, \gamma)$. L_{-b} carries $\triangle ABC$ to $\triangle U_0 B' C'$ where $C' = P(b, \pi)$ and $B' = P(c, \pi - \alpha)$. Together with (6.4) and (6.9), the equation $[B'] = [L_{-b}][B]$ (cf. (5.3)) yields

$$\begin{bmatrix} \cosh c \\ -\sinh c \cos \alpha \\ \sinh c \sin \alpha \end{bmatrix} = \begin{bmatrix} \cosh b & -\sinh b & 0 \\ -\sinh b & \cosh b & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cosh a \\ \sinh a \cos \gamma \\ \sinh a \sin \gamma \end{bmatrix}. \quad (9.1)$$

After expanding, the top entry gives the hyperbolic law of cosines and the bottom gives the law of sines. These imply the rest of trigonometry, including a formula for the angle of parallelism [6, Chapter VI], [8, Chapter 10].

10. EQUATIONS AND POLES OF LINES. Each H -line l is the intersection of H^2 with a subspace of M^3 with equation

$$p(V, X) = -v_0 x_0 + v_1 x_1 + v_2 x_2 = 0 \quad (10.1)$$

for some nonzero $V = \sum v_i U_i \in M^3$. We can regard (10.1) as an equation of l ; set $l = l[V]$. It is not hard to see that $q(V) > 0$; dividing by $\sqrt{q(V)}$, we can assume that V is in the subset

$$D^2 = \{V \in M^3 | q(V) = -v_0^2 + v_1^2 + v_2^2 = 1\} \quad (10.2)$$

of M^3 , a hyperboloid of one sheet. Each line corresponds to exactly two points $\pm V$ of D^2 ; by the analogy with S^2 , which is becoming a bit strained, we can call these the *poles* of the line. D^2 , or rather the space of which it is a model, has been called the *exterior-hyperbolic plane* [3]; the symbol comes from the name of the cosmologist de Sitter, who studied its analogue $D^4 \subset M^5$ [9, pp. 124–131], [21, p. 745].

11. CYCLES. In analogy with the “small circles” of S^2 , consider the curves C on H^2 that are the intersections of H^2 with planes of M^3 that do not contain O ; call such curves the *cycles* (or *generalized circles*) of H^2 . Any three noncollinear points of H^2 are clearly contained in exactly one cycle. These planes have equations of form

$$p(V, X) = -v_0x_0 + v_1x_1 + v_2x_2 = -k \quad (11.1)$$

for fixed $V \in M^3$ and real k , neither V nor k being zero (cf. (10.1)). There are three cases.

Case I. $q(V) < 0$. Dividing (11.1) by $\pm \sqrt{-q(V)}$ gives $V \in H^2$. There exists $T \in \mathbf{G}$ with $T(V) = U_0$; the curve $C_k = T(C)$, congruent to C , has equation $p(U_0, X) = -k$, or $x_0 = k$. By (3.2), $k > 0$. C_k has polar equation $r = \operatorname{arccosh} k$ (cf. (6.9)), so we call C the *circle* of H^2 with center V and radius $s = \operatorname{arccosh} k$. In M^3 , C is an ellipse or circle. The family of all circles of H^2 with center V and the family of all (concurrent) H -lines through V are orthogonal (cf. Section 7) trajectories, since this is true when $V = U_0$; in this situation, the rotations R_θ of (6.7) fix each of the circles and the family of lines. Since the plane $x_0 = k$ containing C_k is Euclidean by (6.5), the H -arc length of C_k (cf. Section 4) equals its circumference in E^3 , identified with M^3 . By (6.9), its radius in E^3 is $\sinh s$, whence its circumference is $2\pi \sinh s$.

Case II. $q(V) > 0$. We can suppose that $V \in D^2$. Choose $T \in \mathbf{G}$ such that $[T(V)] = T([V]) = H^1$. After multiplying (11.1) by -1 if necessary, $T(V) = U_2$. Then $T(C) = C_k$ has equation $x_2 = k$, $k \neq 0$, whence C_k and C are branches of hyperbolas in M^3 . Consider the H -translations L_t along H^1 . For the point $P_s = (\cosh s)U_0 + (\sinh s)U_2$ of the H -line $x_1 = 0$ coordinatized by H -length,

$$L_t(P_s) = (\cosh s \cosh t)U_0 + (\cosh s \sinh t)U_1 + (\sinh s)U_2. \quad (11.2)$$

Therefore L_t maps C_k on itself and $k = \pm \sinh s$ where s is the perpendicular H -distance from any point $L_t(P_s)$ of C_k to H^1 . Accordingly each curve $x_2 = k$ (or $x_2 = \pm k$) is called an *equidistant curve* with axis H^1 . The images under the mappings L_t of the line $x_1 = 0$ form the family of “divergent” or “hyperparallel” H -lines perpendicular to H^1 ; this and the family of all curves C_k are orthogonal trajectories. For fixed $s > 0$, the H -arc length of C_k from P_s to $L_t(P_s)$ is $\int_0^t \sqrt{q(L_u(P_s))} du = t \cosh s$, t being the H -length of the segment of H^1 from U_0 to $L_t(U_0)$.

Case III. $q(V) = 0$. V is on the cone $v_0^2 = v_1^2 + v_2^2$. We can suppose that $v_0 = 1$, so that C has equation $-x_0 + (\cos \theta)x_1 + (\sin \theta)x_2 = -k$. Then $R_{-\theta}(C) = C_k$ has equation $-x_0 + x_1 = -k$. Since C_k lies on H^2 , $k > 0$. The curves C are called *horocycles*; they are parabolas in M^3 . A matrix calculation, as in Section 6, shows that the elements of \mathbf{G} that map any one of the curves C_k on itself are precisely the transformations $N_t J_t^j$ where

$$N_t = \begin{bmatrix} 1 + \frac{t^2}{2} & -\frac{t^2}{2} & t \\ \frac{t^2}{2} & 1 - \frac{t^2}{2} & t \\ t & -t & 1 \end{bmatrix} \quad (11.3)$$

(see [27, p. 172]); here $N_t N_u = N_{t+u}$. Easily,

$$N_t(L_s(U_0)) = \left(\cosh s + \frac{t^2}{2} e^{-s} \right) U_0 + \left(\sinh s + \frac{t^2}{2} e^{-s} \right) U_1 + (te^{-s}) U_2. \quad (11.4)$$

For fixed s this parametrizes the horocycle C_k for which $k = \cosh s - \sinh s = e^{-s}$. The H -arc length on C_k between $L_s(U_0)$ and $N_t(L_s(U_0))$ is $|te^{-s}|$ by a very easy integration.

The H -line $N_t(H^1)$ has equation $tx_0 + tx_1 + x_2 = 0$ and is parametrized by $N_t(L_s(U_0))$ with t fixed; these lines form a family of “parallel” (in the asymptotic sense) lines, the orthogonal trajectories of the curves C_k . The distance between C_1 and C_k along any of these lines, e.g. H^1 , is $|s|$. In this case there is an extra type of symmetry: L_s maps C_1 on C_k , so that all horocycles are congruent. Accordingly, the L_s as well as the N_t fix the two orthogonal families; note that $N_t L_s = L_s N_{te^{-s}}$.

12. AREAS. It is easy to assign an area to each well-behaved region R in H^2 in such a way that areas will be *invariant under congruence* and *additive* [22, p. 154]; define the H -area of R to be any constant times the volume in M^3 of the solid consisting of all points of all segments in M^3 joining O to points of R . By Section 8, congruence amounts to applying some $T \in G$; T preserves volumes since it has determinant ± 1 by Section 5, hence T preserves H -areas and is clearly additive. (The natural choice of the constant is 3.) A similar argument interprets H -length in terms of the area of a sector of a hyperbola (this goes back to [14, p. 260]) and corresponds to a well-known interpretation of hyperbolic functions (see many calculus books and [4, p. 253].)

13. HYPERBOLIC CARTOGRAPHY. Despite the special place that the hyperboloidal model H^2 holds, it is good to have a variety of plane models, in their proper roles as map projections, to bring out different aspects of hyperbolic geometry. The analogy with map projections of the globe [2], [18] is not new, but it unifies many of the known models (see [6, pp. 136–137] and [8, Chapter 7] for example) and encourages the development of new ones. Here are a few examples, without proofs.

An important class consists of the “azimuthal” projections, in which the point $P(r, \theta)$ of H^2 (see (6.9)) is mapped on the point of the Euclidean plane with polar coordinates $(f(r), \theta)$ for some function f . (Terms used by cartographers for the sphere are in quotation marks.) The “gnomonic” projection with $f(r) = \tanh r = (\sinh r / \cosh r)$ gives the projective or Beltrami-Klein model on the unit disk, which represents H -lines by line-segments. The “stereographic” projection with $f(r) = 2 \tanh(r/2)$ gives the Poincaré disk model, which represents H -lines by circles; this is *conformal* in the sense that it preserves angles between curves. These two models are among the best-known, and their connection with cartography has been pointed out by Coxeter [4, pp. 255 and 258] and Penrose [23, p. 113]. (Milnor [19] also discusses these models.) Some other choices are $f(r) = 2 \sinh(r/2)$, which preserves areas; $f(r) = \sinh r$, the “orthographic” projection, which projects H^2 perpendicularly on the Euclidean plane $x_0 = 0$ (see Gans [7]); and of course the “equidistant” projection $f(r) = r$, whose spherical counterpart appears on the seal of the United Nations.

An example that I have not seen worked out is the analogue of the Mercator projection, a famous conformal map of the sphere with “parallels of latitude” and “meridians” represented by orthogonal families of lines. (This analogue, however,

closely resembles the conformal map of D^2 given on [9, pp. 126–127].) With H^1 in the role of the “equator”, this analogue maps the point $L_t(P_s)$ of (11.2) on the point with rectangular coordinates $(t, f(s))$, where f is chosen so that the map is conformal. Since the shapes of infinitesimal regions are preserved, the formula for arc length on the equidistant curves corresponding to parallels of latitude implies that $f'(s) = 1/\cosh s$, whence $f(s) = \arctan(\sinh s)$, the so-called gudermannian function. FIGURE 2 shows the grid of this projection, with H^1 horizontal and with spacings $\frac{1}{4}$. The horizontal lines (extended infinitely on both sides) represent these equidistant curves and H^1 ; the vertical line-segments, of length π , represent the meridians, which are divergent H -lines. The top and bottom lines do not correspond to points of H^2 . We cannot enliven the picture by drawing outlines of continents, but FIGURE 3 shows the image of a polar grid of points $P(r, \theta)$ (“transverse Mercator projection”). FIGURE 4, rotated 90 degrees, represents the image of a tessellation of H^2 by congruent triangles; this is the tessellation used, in the stereographic projection, by M. C. Escher in his picture “Circle Limit IV” [5].

In fact the above is only one of *three* analogues of the Mercator projection, each conformal with a family of cycles playing the role of parallels and with the orthogonal trajectory lines as meridians. While the projection based on circles seems of little interest, the one based on horocycles gives the Poincaré half-plane model with the horocycles represented by parallel lines and their orthogonal trajectories by half-lines. This model, usually taken on the upper half of the

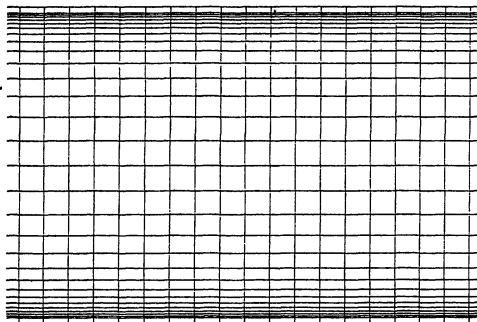


Figure 2

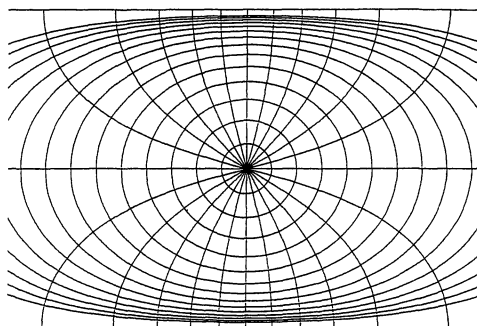


Figure 3

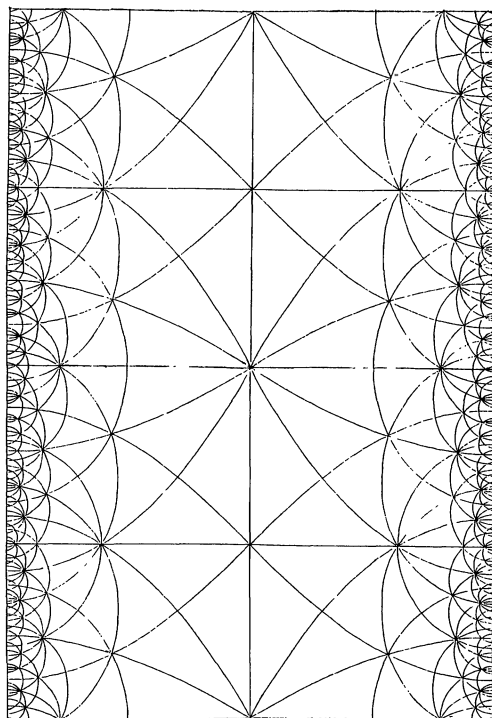


Figure 4

complex number plane, is extremely important because of its wide-ranging applications (e.g. [16], [17].)

14. ON THE HISTORY OF THE MODEL. The high points of the prehistory of the model H^2 are Lambert's idea of an "imaginary sphere" in 1766 [1, p. 50] and Taurinus' calculation of trigonometry on a "sphere of imaginary radius" in 1826 [1, p. 79]. The name "hyperbolic geometry", introduced by Klein in 1871, does not refer to this model; see [31, p. 63].

Clear statements of H^2 seem to have originated twice. According to Killing's 1885 book [14, pp. 258–259], Weierstrass communicated the coordinates x_i at a seminar in 1872 and "supplied numerous applications". Accordingly Killing named them "Weierstrass coordinates" and used them extensively. (Killing remarked that Beltrami had previously used some closely related imaginary coordinates, but only in one proof.) There is a distinction, however, between using the coordinates and using the points of a hyperboloid as a model. Killing made it clear that he was doing the latter in a one-paragraph presentation [14, p. 260] of H^2 as an "Abbildung" (mapping, image, picture) of the hyperbolic plane in E^3 , which included a description of H -lines and cycles. In 1880, he had described [13, p. 275] the exterior-hyperbolic plane as an ideal region of the hyperbolic plane; although he used Weierstrass coordinates there, he mapped the hyperbolic plane on a hemisphere instead of a hyperboloid [13, pp. 284–287]. See Hawkins [10] for a real historian's treatment of Killing.

Meanwhile, in 1881, Poincaré [24] took quantities corresponding to our x_i from an 1854 paper of Hermite on quadratic forms, named them "hyperbolic coordi-

nates” without mentioning a hyperboloid, and related them to the projective disk model. (Some unpublished work of Poincaré in 1880 related to our model is discussed by Gray [7a, pp. 271–272].) In 1887, he defined [25] “quadratic geometries” on arbitrary quadric surfaces in 3-space, with distances and angles defined in terms of cross-ratios. These included spherical, hyperbolic (a form of our model), exterior-hyperbolic, and Euclidean geometries, the last on a paraboloid!

The latest part of our model to be developed was the connection with the Minkowski structure of the embedding space. Naturally enough, this seems to have followed the origin of special relativity. Some early references are Minkowski [20, p. 376] (a posthumous paper based on a 1907 lecture), Sommerfeld [28], and Varičák [30]. In 1909, Jansen [12] gave what appears to be the first detailed exposition of H^2 , referring to Poincaré and Minkowski; he derived most properties by translating them from the half-plane model.

The “imaginary radius” idea survived, mixed with H^2 , in Klein’s classic posthumous book [15, p. 193] and more recently in [26].

REFERENCES

1. R. Bonola, *Non-Euclidean Geometry*, Dover, New York, 1955.
2. B. H. Brown, Conformal and equiareal world maps, this MONTHLY, 42 (1935) 212–223. Also in *Selected Papers on Geometry*, The Raymond W. Brink Selected Mathematical Papers, Vol. 4, Math. Assn. of America, Washington, 1979, pp. 29–39.
3. H. S. M. Coxeter, A geometrical background for de Sitter’s world, this MONTHLY, 50 (1943) 217–228.
4. ———, *Non-Euclidean Geometry*, 5th edition, University of Toronto Press, 1965.
5. ———, Angels and devils, in D. A. Klarner (ed.), *The Mathematical Gardner*, Prindle, Weber, and Schmidt, Boston, 1981, pp. 197–210.
6. R. L. Faber, *Foundations of Euclidean and Non-Euclidean Geometry*, Dekker, New York/Basel, 1983.
7. D. Gans, A new model of the hyperbolic plane, this MONTHLY, 73 (1966) 291–295.
- 7a. J. J. Gray, *Linear Differential Equations and Group Theory from Riemann to Poincaré*, Birkhäuser, Boston/Basel, 1986.
8. M. J. Greenberg, *Euclidean and Non-Euclidean Geometries*, 2nd edition, Freeman, San Francisco, 1980.
9. S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, 1973.
10. T. Hawkins, Non-Euclidean geometry and Weierstrassian mathematics: the background to Killing’s work on Lie algebras, *Historia Mathematica*, 7 (1980) 289–342.
11. N. Jacobson, *Basic Algebra I*, 2nd edition, Freeman, San Francisco, 1985.
12. H. Jansen, Abbildung der hyperbolischen Geometrie auf ein zweischaliges Hyperboloid, *Mitt. Math. Gesellsch. Hamburg*, 4 (1909) 409–440.
13. W. Killing, Die Rechnung in den Nicht-Euklidischen Raumformen, *J. Reine Angew. Math.*, 89 (1880), 265–287.
14. ———, *Die Nicht-Euklidischen Raumformen in Analytischer Behandlung*, Teubner, Leipzig, 1885.
15. F. Klein, *Vorlesungen über Nicht-euklidische Geometrie*, reprint, Chelsea, New York, 1959.
16. G. Mason, Finite groups and modular functions (with an appendix by S. P. Norton), *Proc. Symp. Pure Math.*, 47 (1987), vol. 1, 181–210.
17. B. Mazur, Number theory as gadfly, this MONTHLY, 98 (1991) 593–610.
18. J. Milnor, A problem in cartography, this MONTHLY, 76 (1969) 1101–1112. Also in *Selected Papers in Geometry* (see reference 2), pp. 180–191.
19. ———, Hyperbolic geometry: the first 150 years, *Bull. (New Series) Amer. Math. Soc.*, 6 (1982), 9–24.
20. H. Minkowski, Das Relativitätsprinzip, *Jber. Deutsch. Math.-Verein.*, 24 (1915) 372–382. Also *Ann. Physik*, (4) 47 (1915), 927–938.
21. C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, Freeman, San Francisco, 1973.
22. E. E. Moise, *Elementary Geometry from an Advanced Standpoint*, 2nd edition, Addison-Wesley, Reading, Mass., 1974.

23. R. Penrose, The geometry of the universe, in *Mathematics Today: Twelve Informal Essays*, L. A. Steen, ed., Springer, Berlin, 1978, pp. 83–125.
24. H. Poincaré, Sur les applications de la géométrie non euclidienne à la théorie des formes quadratiques, *Compte Rendu de l'Association Française pour l'Avancement des Sciences, 10^e Session 1881* (Alger), pp. 132–138. Also in *Œuvres*, vol. 5, Gauthier-Villars, Paris (1950), pp. 267–274.
25. ———, Sur les hypothèses fondamentales de la géométrie, *Bull. Soc. Math. France*, 15 (1887) 203–216. Also in *Œuvres*, vol. 11, Gauthier-Villars, Paris (1956), pp. 79–91.
26. G. Y. Rainich and S. M. Dowdy, *Geometry for Teachers, An Introduction to Geometrical Theories*, Wiley, New York, 1968.
27. B. A. Rozenfel'd, *Non-Euclidean Spaces* (Russian), Nauka, Moscow, 1969.
28. A. Sommerfeld, Über die Zusammensetzung der Geschwindigkeiten in der Relativtheorie, *Physikalische Zeitschrift*, 10 (1909) 826–829.
29. M. Spivak, *A Comprehensive Introduction to Differential Geometry*, vol. 4, Publish or Perish, Inc., Boston, 1975.
30. V. Varičák, Anwendung der Lobatschefskijschen Geometrie in der Relativtheorie, *Physikalische Zeitschrift*, 11 (1910) 93–96.
31. H. E. Wolfe, *Introduction to Non-Euclidean Geometry*, Holt, Rinehart and Winston, New York, 1945.

Department of Mathematics
Tufts University
Medford, MA 02155
e-mail: reynolds@jade.tufts.edu

Lester R. Ford Award

The Lester R. Ford Award for an
 expository article published in the
 MONTHLY in 1991 was presented at the
 San Antonio joint meetings to

C. W. H. Lam

for his article *The Search for a Finite
 Projective Plane of Order 10*, 98(1991),
 305–318.

A Simple Heuristic Proof of Hardy and Littlewood's Conjecture B

Michael Rubinstein

Hardy and Littlewood have conjectured (see [3]) that *there are infinitely many prime pairs $(p, p + m)$ for every even m . If $\pi_m(x)$ is the number of pairs less than x , then*

$$\pi_m(x) \sim 2C_2 \frac{x}{(\log x)^2} \prod_{\substack{p>2 \\ p|m}} \frac{p-1}{p-2} \quad (*)$$

where $C_2 = \prod_{p>2} (1 - 1/(p-1)^2)$.

If $m = 2$ then we call pairs of primes $(p, p + 2)$ twin primes. Thus, in particular, (*) implies that

$$\pi_2(x) \sim 2C_2 \frac{x}{(\log x)^2}. \quad (**)$$

In this paper, I present a short heuristic proof of (**) and generalize to give a heuristic proof of (*).

For other heuristic proofs consult [2] and [4].

“Proof” We begin by considering the following combinatorial problem: Let $A = \{1, 2, 3, \dots, a\}$ and choose from A a subset B consisting of b elements. Now, say we choose another subset B' from A consisting of b' elements. The elements in B' are chosen randomly (by randomly I mean that every element of A has an equal chance of being picked). Then, the expected number of elements in $B \cap B'$ is equal to $b' \cdot (b/a) = b'b/a$. Since this trivial result is used throughout this article, we refer to it as (1).

We also reference Dirichlet's Theorem (see [1] for a proof of Dirichlet's Theorem) which states that given an arithmetical progression $ak + b$, where $(a, b) = 1$, then

$$\sum_{\substack{p \in ak+b \\ p \leq x}} 1 \sim \frac{x}{\phi(a) \log x}. \quad (2)$$

We now give a heuristic proof of (**).

Consider the following 2 pairs of arithmetical progressions:

$$(2k + 0), (2k + 2)$$

$$(2k + 1), (2k + 3)$$

where $k = 0, 1, 2, 3 \dots$

All twin primes will fall in the second pair $(2k + 1), (2k + 3)$. Therefore, the question of how many twin primes exist up to a certain x is equivalent to finding out how many values of k make $2k + 1$ and $2k + 3$ simultaneously prime.

Now, assume that any value of k is equally likely to make $2k + 1$ prime, and any value of k is equally likely to make $2k + 3$ prime (this assumption is incorrect and will be improved on shortly). Further, assume that the primes belonging to $2k + 1$ behave independently from the primes belonging to $2k + 3$. Thus, using (1) and (2), we have that

$$\pi_2(x) \sim \frac{\left(\frac{x}{\phi(2)\log x}\right)^2}{\frac{x}{2}} = 2 \frac{x}{(\log x)^2}.$$

In the above,

$$b \sim b' \sim \frac{x}{\phi(2)\log x}, \quad \text{and} \quad a = \left\lfloor \frac{x}{2} \right\rfloor \sim \frac{x}{2}.$$

The $x/(\log x)^2$ part reflects the conjecture, but the constant 2 does not. The assumption that any value of k is equally likely to make $2k + 1$ prime, and that any value of k is equally likely to make $2k + 3$ prime is incorrect. For, if $k \equiv 1 \pmod 3$, then $2k + 1$ is composite ($k > 1$), and if $k \equiv 0 \pmod 3$, then $2k + 3$ is composite ($k > 0$). And so, we improve on the above result by considering pairs of arithmetical progressions mod 6 rather than pairs of arithmetical progressions mod 2. Thus, consider the following 6 pairs of arithmetical progressions:

$$(6k + 0), (6k + 2)$$

$$(6k + 1), (6k + 3)$$

$$(6k + 2), (6k + 4)$$

$$(6k + 3), (6k + 5)$$

$$(6k + 4), (6k + 6)$$

$$(6k + 5), (6k + 7)$$

We can eliminate all pairs except for $(6k + 5), (6k + 7)$ as the others cannot contribute anything to $\pi_2(x)$ (except when $k = 0$). Thus by (1) and (2), we have that,

$$\pi_2(x) \sim \frac{\left(\frac{x}{\phi(2 \cdot 3)\log x}\right)^2}{\frac{x}{2 \cdot 3}} = \frac{3}{2} \frac{x}{(\log x)^2}.$$

Somewhat better.

However, once again not every value of k is equally likely to make $6k + 5$ prime, and not every value of k is equally likely to make $6k + 7$ prime. For, if $k \equiv 0 \pmod 5$, then $6k + 5$ is composite ($k > 0$), and if $k \equiv 3 \pmod 5$ then $6k + 7$ is composite. And so, we improve on the above by considering pairs of arithmetical progressions mod 30 instead of pairs of arithmetical progressions mod 6. Thus, consider the following 30 pairs of arithmetical progressions:

$$(30k + 0), (30k + 2)$$

$$(30k + 1), (30k + 3)$$

$$(30k + 29), (30k + 31)$$

The only three pairs that contribute to $\pi_2(x)$ are $(30k + 11), (30k + 13), (30k + 17), (30k + 19), (30k + 29), (30k + 31)$ (with the obvious few finite exceptions such as 3, 5 when $k = 0$).

Now, by (1) and (2), each pair contributes

$$\frac{\left(\frac{x}{\phi(2 \cdot 3 \cdot 5) \log x}\right)^2}{\frac{x}{2 \cdot 3 \cdot 5}} = \frac{2 \cdot 3 \cdot 5 x}{(1 \cdot 2 \cdot 4 \log x)^2} = \frac{15}{32} \frac{x}{(\log x)^2}$$

to $\pi_2(x)$. Thus

$$\pi_2(x) \sim 3 \left(\frac{15}{32} \frac{x}{(\log x)^2} \right).$$

Once again, not all values of k are equally likely to make $30k + b$ prime. And so we continue in the above fashion indefinitely, and find that

$$\begin{aligned} \pi_2(x) &\sim \lim_{k \rightarrow \infty} \phi_2(2 \cdot 3 \cdot 5 \cdots p_k) \left(\frac{\left(\frac{x}{\phi(2 \cdot 3 \cdot 5 \cdots p_k) \log x}\right)^2}{\frac{x}{2 \cdot 3 \cdot 5 \cdots p_k}} \right) \\ &= \lim_{k \rightarrow \infty} \frac{\phi_2(2 \cdot 3 \cdot 5 \cdots p_k)(2 \cdot 3 \cdot 5 \cdots p_k)}{(\phi(2 \cdot 3 \cdot 5 \cdots p_k))^2} \cdot \frac{x}{(\log x)^2}, \end{aligned}$$

where $\phi_2(n)$ denotes the number of pairs $c, c + 2$ such that $(n, c) = (n, c + 2) = 1$, and $0 \leq c \leq n - 1$ (i.e. the number of pairs of arithmetical progressions $(nk + c), (nk + c + 2)$ that contain infinitely many pairs of primes).

The reader will object that we are unjustified in passing to the limit since (2) is no longer valid when a (in (2)) is infinite. We would have to fix this problem by letting k (in the above) be a function of x (so that $p_k \ll x$), and use a more precise version of (2). However, since the method being described is heuristic, we omit the details, and proceed.

I claim that $\phi_2(2 \cdot 3 \cdot 5 \cdot 7 \cdots p_k) = (3 - 2)(5 - 2)(7 - 2)(11 - 2) \cdots (p_k - 2)$. This will be proven shortly. Assuming this for the moment, we see that the above is simply

$$\pi_2(x) \sim 2 \prod_{p > 2} \frac{(p - 2)p}{(p - 1)^2} \frac{x}{(\log x)^2} = 2 \prod_{p > 2} \left(1 - \frac{1}{(p - 1)^2} \right) \frac{x}{(\log x)^2}.$$

We have arrived at (**).

It remains to be shown that $\phi_2(2 \cdot 3 \cdot 5 \cdots p_k) = (3 - 2)(5 - 2) \cdots (p_k - 2)$. In fact, I prove that.

Lemma 1. *Given $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_j^{\alpha_j}$ where the p_i 's are arbitrary primes and $\alpha_i \geq 1$, then*

$$\phi_2(n) = p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} \cdots p_j^{\alpha_j - 1} \prod_{p_i > 2} (p_i - 2).$$

Note the similarity to Euler's totient function.

Proof: Two steps are required:

a) Show that it is true for p^α , p a prime. That,

$$\begin{aligned}\phi_2(p^\alpha) &= p^{\alpha-1} \cdot (p-2) && (\text{if } p > 2) \\ &= p^{\alpha-1} && (\text{if } p = 2)\end{aligned}$$

b) Show that $\phi_2(ab) = \phi_2(a)\phi_2(b)$ if $(a, b) = 1$.

Proof of a): If $p > 2$ then, of the integers $0, 1, 2, 3, 4, \dots, p^\alpha - 1$ there are exactly $p^{\alpha-1}$ integers that are not relatively prime to p^α (namely $0, p, 2p, 3p, 4p, \dots, (p^{\alpha-1} - 1)p$). Each of these $p^{\alpha-1}$ integers kills two arithmetical progressions and so $\phi_2(p^\alpha) = p^\alpha - 2p^{\alpha-1} = p^{\alpha-1}(p-2)$.

If $p = 2$, then each multiple of p only kills one pair of arithmetical progressions (since we must not double count) and so $\phi_2(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}$ (since $p = 2$).

Proof of b): Partition the integers $0, 1, 2, 3, \dots, ab - 1$ as follows:

$$\begin{array}{cccc} 0, 1, \dots, a-1, & a, \dots, 2a-1, & 2a, \dots, 3a-1, & \dots, (b-1)a, \dots, ba-1. \\ A1 & A2 & A3 & Ab\end{array}$$

Now, in each Ai , there are $\phi_2(a)$ pairs $c, c+2$, where $(i-1)a \leq c \leq ia-1$, such that $(a, c) = (a, c+2) = 1$. Furthermore, each pair appears b times modulo a (i.e. once in each Ai). We need to find how many of these pairs are relatively prime to b . Well, for any pair $c, c+2$, where $0 \leq c \leq a-1$, we may list the b times that it appears modulo a :

$$c, c+2, a+c, a+c+2, \dots, (b-1)a+c, (b-1)a+c+2. \quad (3)$$

If we examine these pairs modulo b , we see that, since $(a, b) = 1$, they run through all pairs $i, i+2 \pmod{b}$. And so, of the pairs in (3), $\phi_2(b)$ are relatively prime to b . Thus, the total number of pairs relatively prime to a and b is equal to $\phi_2(a)\phi_2(b)$.

The above method that I've described above for twin primes may be generalized to work for (*). In fact, an identical argument is used, the only difference being in our counting function $\phi_2(n)$. I leave out the details (as they are more or less identical) but summarize the results below. The reader should fill in the details while taking a shower.

Given m is an even integer then

$$\begin{aligned}\pi_m(x) &\sim \lim_{k \rightarrow \infty} \phi_m(2 \cdot 3 \cdot 5 \cdots p_k) \left(\frac{\left(\frac{x}{\phi(2 \cdot 3 \cdot 5 \cdots p_k) \log x} \right)^2}{\frac{x}{2 \cdot 3 \cdot 5 \cdots p_k}} \right) \\ &= \lim_{k \rightarrow \infty} \frac{\phi_m(2 \cdot 3 \cdot 5 \cdots p_k)(2 \cdot 3 \cdot 5 \cdots p_k)x}{(\phi(2 \cdot 3 \cdot 5 \cdots p_k))^2(\log x)^2},\end{aligned}$$

where $\phi_m(n)$ denotes the number of pairs $c, c+m$ such that $(n, c) = (n, c+m) = 1$, and $0 \leq c \leq n-1$ (i.e. the number of pairs of arithmetical progressions $(nk+c), (nk+c+m)$ that contain infinitely many primes).

In a similar manner to the proof of lemma 1, it is easy to establish that

$$\begin{aligned}\phi_m(p^\alpha) &= p^{\alpha-1} \cdot (p-2) && (\text{if } (p, m) = 1, p \text{ a prime}) \\ &= p^{\alpha-1}(p-1) && (\text{if } p|m, p \text{ a prime})\end{aligned}$$

and that

$$\phi_m(ab) = \phi_m(a)\phi_m(b) \quad \text{if } (a, b) = 1.$$

Thus, in particular we have,

$$\begin{aligned}\phi_m(2 \cdot 3 \cdot 5 \cdots p_k) &= \prod_{\substack{p \leq p_k \\ (p, m) = 1}} (p-2) \prod_{\substack{p \leq p_k \\ p|m}} (p-1) \\ &= \prod_{2 < p \leq p_k} (p-2) \prod_{\substack{2 < p \leq p_k \\ p|m}} \frac{p-1}{p-2}.\end{aligned}$$

And so,

$$\begin{aligned}\pi_m(x) &\sim 2 \left(\prod_{p>2} \frac{(p-2)p}{(p-1)^2} \right) \left(\prod_{\substack{p>2 \\ p|m}} \frac{p-1}{p-2} \right) \frac{x}{(\log x)^2} \\ &= 2C_2 \frac{x}{(\log x)^2} \prod_{\substack{p>2 \\ p|m}} \frac{p-1}{p-2}.\end{aligned}$$

We have arrived at (*).

CONCLUSION. The beauty of the heuristic method that I have described in this article lies in its simplicity, and in the fact that it gives the conjectured value. It presents the reader with a somewhat convincing reason as to why the conjecture ought to be true. It should also be remarked that the same strategy can be applied to similar problems (such as pairs of primes that differ linearly and not just by a constant).

REFERENCES

1. Apostol, *An Introduction to Analytic Number Theory*, 1976, Springer-Verlag, New York, 146–156.
2. Cherwell, *Quarterly Journal of Mathematics* (Oxford), 17 (1946) 46–62.
3. Hardy and Littlewood, *Acta Math.*, 44 (1923) 1–70.
4. Polya, *American Math. Monthly*, 66 (1959) 375–384.

68 Banstead Rd.
Montreal West, Quebec
Canada, H4X 1P2
miker@phoenix.princeton.edu

The Pompeiu Problem

H. Turner Laquer

1. INTRODUCTION. The classical Pompeiu problem asks if a continuous function of two variables must be identically zero if the integrals of the function over all disks of a fixed radius are zero. In this paper we will be interested in the following similar problem:

Problem. *Is it the case that a continuous function on the 2-sphere must be identically zero if the integrals of the function over all α -sectors are zero?*

By definition, an α -sector is a region between two great circles intersecting at an angle α , i.e. a spherical bi-angle with both angles equal to α (see FIGURE 1).

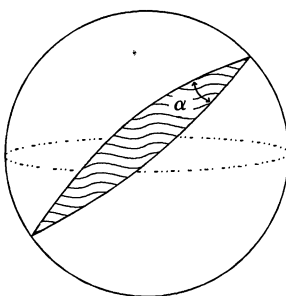


Figure 1. An α -sector on the 2-sphere.

The main objective of this paper is to give a positive solution to the problem under the assumption that the function is at least \mathcal{C}^1 (see Theorem 1).

The solution to the problem is essentially elementary, using little more than multivariable calculus. Surprisingly, the solution makes use of all four of the classical integrals typically covered in a course on multivariable calculus, namely, the scalar surface, surface flux, line, and path integrals. In addition, there is an appealing application of Stokes' theorem resulting from the development of certain perturbation formulas (see sections 3 and 4). The proof of the main result is given in section 5. Finally, section 6 contains a computational lemma which is the only part of the main argument which seems to require some limited use of representation theory.

2. POMPEIU PROBLEMS. Historically, questions of the type described in the introduction have been associated with the name Pompeiu. The most classical of such problems deals with integrals over disks of a fixed radius of a function of two

variables. Although this problem has apparently had an interesting history (see [2, 4, 9, 10, 11]), it is a relatively straightforward calculus problem to show that for any given radius ρ there are functions of the form $f(x, y) = \sin(ax + by)$, for appropriate a and b , all of whose integrals over disks of radius ρ are zero. The proof is not completely elementary in that the actual integrals involve the first order Bessel function [1].

A problem equivalent to the classical Pompeiu problem has been studied in the case of the 2-sphere. This goes as follows. Let $D_\rho(\vec{p})$ be the spherical cap of radius ρ centered at a point \vec{p} on the 2-sphere and let $T_\rho: \mathfrak{F}(S^2) \rightarrow \mathfrak{F}(S^2)$ be the integral transformation

$$(T_\rho f)(\vec{p}) = \iint_{D_\rho(\vec{p})} f dA.$$

Ungar's "Freak Theorem" [8] states that: the set of radii for which T_ρ is not injective forms a countable dense subset of the interval $[0, \pi]$. So for a countable dense subset of radii, there exist nonzero smooth functions on the 2-sphere all of whose integrals over spherical disks of the given fixed radius are zero while for all other radii there are no such functions.

Ungar's theorem and the classical Pompeiu problem have a variety of generalizations. One possibility is to consider integrals over "disks" in other spaces. Another possibility is to allow more general regions of integration. These regions should still be of the same dimension as the original space, i.e. open subsets with reasonable boundaries. This is in contrast to Radon-type transformations which involve integrals over lower dimensional subspaces [7].

If D is a region in S^2 , then integrals over translates of D define an integral transformation $T_D: \mathfrak{F}(S^2) \rightarrow \mathfrak{F}(SO(3))$ by letting

$$(T_D f)(g) = \iint_{g \cdot D} f dA \quad \forall g \in SO(3).$$

Due to the lack of symmetry of a general region, it is necessary that the new function $T_D f$ be defined on $SO(3)$ —the group of rotations of the sphere. The "generalized Pompeiu problem" is to describe those regions which yield an injective transformation.

Naively, it might appear that T_D would be injective whenever the region D is not symmetric. This hope stems from the three dimensional nature of the set of rigid motions of the sphere. However, this is not the case. Noninjective spherical disks can be combined to create certain obvious noninjective regions. Such examples can be eliminated by requiring that the boundary of D be homeomorphic to S^1 . However, this is still not sufficient to ensure injectivity. Ungar's paper already includes a sketch of an existence proof for a spherical polygon (edges being parts of great circles) which results in a noninjective transformation [8]. It is unclear how many sides would be required for such a polygon. The following theorem, however, shows that such a polygon must necessarily have at least 3 edges.

Theorem 1. *If the integrals of a \mathcal{C}^1 -function on the 2-sphere over all α -sectors are zero, then the function is identically zero (α fixed with $0 < \alpha < \pi$).*

The proof of Theorem 1 will be given in section 5.

3. THE PRIMARY PERTURBATION. The main idea used in the proof of Theorem 1 is “perturbation”. Suppose D_0 is a reasonable region (surface) in S^2 and let f be some function on S^2 . If the region D_0 is moved slightly then the change in the integral of f can only depend on the values of f along the boundary of D_0 . The values of f in the interior of D_0 will not contribute to the change. This indicates that it must be possible to apply Stokes’ theorem! More precisely, there is the following theorem: (see FIGURE 2).

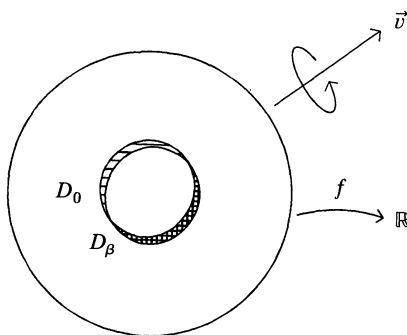


Figure 2. Theorem 2 shows that when the region D_0 is moved slightly, by means of a rotation around the axis \vec{v} , all the change in the integral of f is given by a line integral along the boundary of D_0 of the vector field $f\vec{v}$. Note how parts of the boundary perpendicular to \vec{v} give no contribution (to first order) while the contribution is greatest (appropriately signed) when the boundary and \vec{v} are parallel.

Theorem 2. Let D_0 be a region in the unit sphere to which Stokes’ theorem can be applied and let D_β be the region D_0 rotated by an angle β around an axis determined by a unit vector \vec{v} . Then for any \mathcal{C}^1 -function f on S^2 ,

$$\left. \frac{d}{d\beta} \right|_{\beta=0} \iint_{D_\beta} f dA = \int_{\partial D_0} f \vec{v} \cdot d\vec{s}.$$

Proof: Let $L_\beta(\vec{v})$ be the rotation around the axis determined by \vec{v} . Vector geometry gives

$$L_\beta(\vec{v})(\vec{w}) = (1 - \cos \beta)(\vec{v} \cdot \vec{w})\vec{v} + (\cos \beta)\vec{w} + (\sin \beta)(\vec{v} \times \vec{w}).$$

Change of variables and differentiation under the integral yield

$$\left. \frac{d}{d\beta} \right|_{\beta=0} \iint_{D_\beta} f dA = \iint_{D_0} \left. \frac{\partial}{\partial \beta} \right|_{\beta=0} (f \circ L_\beta(\vec{v})) dA.$$

Note: this also uses the fact that the area element of the unit sphere is rotationally invariant. By the chain rule

$$\begin{aligned} \left. \frac{\partial}{\partial \beta} \right|_{\beta=0} (f \circ L_\beta(\vec{v}))(\vec{w}) &= \nabla f(L_0(\vec{v})(\vec{w})) \cdot \left(\left. \frac{\partial}{\partial \beta} \right|_{\beta=0} L_\beta(\vec{v})(\vec{w}) \right) \\ &= \nabla f(\vec{w}) \cdot (\vec{v} \times \vec{w}). \end{aligned}$$

Standard formulas in vector analysis [6] show that

$$\begin{aligned} \nabla f(\vec{w}) \cdot (\vec{v} \times \vec{w}) &= (\nabla f(\vec{w}) \times \vec{v}) \cdot \vec{w} \\ &= ((\nabla \times f\vec{v})(\vec{w})) \cdot \vec{w}. \end{aligned}$$

Since the unit normal at a point \vec{w} on the unit sphere is the vector \vec{w} , this gives

$$\left. \frac{d}{d\beta} \right|_{\beta=0} \iint_{D_\beta} f dA = \iint_{D_0} (\nabla \times (f\vec{v})) \cdot d\vec{A}.$$

Finally, by Stokes' theorem this expression equals

$$\int_{\partial D_0} f\vec{v} \cdot d\vec{s}. \quad \square$$

Corollary. Let D_0 be the standard α -sector $\{0 \leq \varphi \leq \pi, 0 \leq \theta \leq \alpha\}$ in spherical coordinates. Let $\vec{v} = (\sin \alpha, -\cos \alpha, 0)$ and let D_β and f be as in Theorem 2. Then

$$\left. \frac{d}{d\beta} \right|_{\beta=0} \iint_{D_\beta} f dA = \int_{\varphi=0}^{\pi} f(\sin \varphi, 0, \cos \varphi) \sin \alpha \cos \varphi d\varphi.$$

Proof: The boundary of D_0 consists of two great half circles. Because of the choice of \vec{v} , the line integral of $(f\vec{v})$ along the circle at $\theta = \alpha$ is zero. Thus

$$\begin{aligned} \int_{\partial D_0} (f\vec{v}) \cdot d\vec{s} &= \int_{\varphi=0}^{\pi} f(\sin \varphi, 0, \cos \varphi) (\sin \alpha, -\cos \alpha, 0) \cdot (\cos \varphi, 0, -\sin \varphi) d\varphi \\ &= \int_{\varphi=0}^{\pi} f(\sin \varphi, 0, \cos \varphi) \sin \alpha \cos \varphi d\varphi. \end{aligned} \quad \square$$

4. WEIGHTED PATH INTEGRALS AND THE SECONDARY PERTURBATION.

The integral in the corollary to Theorem 2 is a sort of *weighted path integral* of the function f . More generally, if $\vec{c}(s)$ is any curve parametrized by arc length and if $w(s)$ is some weighting function, then there is the weighted path integral

$$\int_0^L f(\vec{c}(s)) w(s) ds.$$

If the curve $\vec{c}(s)$ is piecewise \mathcal{C}^1 , then there is a natural perturbation of the curve along its length, namely, $\vec{c}_t(s) = \vec{c}(s + t)$ where $\vec{c}(s)$ is extended in a \mathcal{C}^1 -manner for $s < 0$ and $s > L$.

Theorem 3. Let \vec{c} and \vec{c}_t be as above. Then

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \int_0^L f(\vec{c}_t(s)) w(s) ds \\ = f(\vec{c}(L)) w(L) - f(\vec{c}(0)) w(0) - \int_0^L f(\vec{c}(s)) w'(s) ds. \end{aligned}$$

Proof: Differentiation under the integral and the chain rule give

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \int_0^L f(\vec{c}_t(s)) w(s) ds \\ = \int_0^L \nabla f(\vec{c}_0(s)) \cdot (\vec{c}_0(s))' w(s) ds \\ = \int_0^L (f(\vec{c}(s)) w(s))' - f(\vec{c}(s)) w'(s) ds \\ = f(\vec{c}(L)) w(L) - f(\vec{c}(0)) w(0) - \int_0^L f(\vec{c}(s)) w'(s) ds. \end{aligned} \quad \square$$

Unlike the result in Theorem 2, the perturbation in Theorem 3 does not reduce the integration to the boundary, i.e. endpoints, of the curve. This is due to the nonuniform weighting factor in the path integral. The perturbation ideas in Theorems 2 and 3 are clearly quite general. For example, results in [5] show how these ideas can be generalized to a wide class of Pompeiu transformations on homogeneous spaces.

5. THE INJECTIVITY OF THE SECTOR TRANSFORMATION. This section contains the proof of Theorem 1. Suppose f is in the kernel of T_D where D is the standard α -sector $\{0 \leq \varphi \leq \pi, 0 \leq \theta \leq \alpha\}$. The corollary to Theorem 2 shows that

$$\int_0^\pi f(\vec{c}(\varphi)) \cos \varphi \, d\varphi = 0$$

where $\vec{c}(\varphi)$ is any great half circle on the sphere, parametrized by arc length. By the secondary perturbation (Theorem 3) it follows that

$$f(\vec{p}) + f(-\vec{p}) = \int_0^\pi f(\vec{c}(\varphi)) \sin \varphi \, d\varphi$$

for any great half circle $\vec{c}(\varphi)$ joining \vec{p} to $-\vec{p}$.

Since the set of points opposite an α -sector is again an α -sector, f being in the kernel of T_D is equivalent to the even and odd parts of f being in the kernel. This allows consideration of even and odd cases separately.

If f is even, choose a point \vec{p}_0 where f is a maximum. Then

$$\begin{aligned} 2f_{\text{MAX}} &= f(\vec{p}_0) + f(-\vec{p}_0) = \int_0^\pi f(\vec{c}(\varphi)) \sin \varphi \, d\varphi \\ &\leq \int_0^\pi f_{\text{MAX}} \sin \varphi \, d\varphi = 2f_{\text{MAX}}. \end{aligned}$$

Equality can hold only if $f(\vec{c}(\varphi))$ equals f_{MAX} along all great half circles joining \vec{p}_0 to $-\vec{p}_0$. Thus f is constant, hence zero.

When the function f is odd, the secondary perturbation gives

$$0 = \int_0^\pi f(\vec{c}(\varphi)) \sin \varphi \, d\varphi.$$

By integrating this over all half circles from \vec{p}_0 to $-\vec{p}_0$, with the half circles also lying in a hemisphere H of S^2 , it follows that

$$\iint_H f \, dA = \int_{\theta=0}^\pi \int_{\varphi=0}^\pi f(\vec{c}_\theta(\varphi)) \sin \varphi \, d\varphi \, d\theta = 0.$$

The $\sin \varphi$ factor is precisely what is needed for the area element of spherical coordinates. Since this integral must be zero for arbitrary hemispheres, the lemma in section 6 shows that f must be identically zero.

6. A COMPUTATIONAL LEMMA. What follows is the one part in the proof of Theorem 1 which seems to require the use of special functions and representation theory (see [3] for details).

Lemma. *If the integrals of a continuous odd function on the 2-sphere over all hemispheres are zero then the function itself must be identically zero.*

Proof: Let $T: \mathfrak{F}(S^2) \rightarrow \mathfrak{F}(S^2)$ be the integral transformation

$$(Tf)(\vec{p}) = \iint_{H(\vec{p})} f dA$$

where $H(\vec{p})$ is the hemisphere centered at \vec{p} . The left action of the Lie group $SO(3)$ on the 2-sphere defines an action of $SO(3)$ on the space of functions on S^2 by $g \cdot f = f \circ L_{g^{-1}}$. Since this action commutes with the transformation T , T is an *intertwining operator*. Next, the space $\mathfrak{F}(S^2)$ of functions splits into a direct sum $\bigoplus_{n=0}^{\infty} V_n$ of subspaces which are invariant under the action of $SO(3)$. These V_n are inequivalent irreducible $SO(3)$ -representations. By Schur's lemma, the operator T acts as a scalar, say λ_n , on each irreducible piece. A direct computation can determine the values of these scalars.

Each V_n includes a “zonal harmonic” function which, in spherical coordinates, is given by

$$f_n(\varphi, \theta) = P_n(\cos \varphi).$$

The function $P_n(t)$ is the n th Legendre polynomial

$$P_n(t) = \frac{1}{2^n n!} \left(\frac{d}{dt} \right)^n (t^2 - 1)^n.$$

Since $f_n(0, \theta) = P_n(1) = 1$, the constant λ_n is given by the integral of f_n over the upper hemisphere. Specifically,

$$\begin{aligned} \lambda_n &= \int_{\varphi=0}^{\pi/2} \int_{\theta=0}^{2\pi} P_n(\cos \varphi) \sin \varphi \, d\theta \, d\varphi \\ &= 2\pi \int_0^1 P_n(t) \, dt \\ &= \frac{2\pi}{2^n n!} \left(\frac{d}{dt} \right)^{n-1} (t^2 - 1)^n \Big|_{t=0} \\ &= \begin{cases} 2\pi & \text{if } n = 0; \\ 0 & \text{if } n = 2k > 0; \\ 2\pi \cdot \frac{(-1)^k}{2^n \cdot n} \cdot \binom{n}{k} & \text{if } n = 2k + 1. \end{cases} \end{aligned}$$

The lemma now follows because the odd functions on the sphere correspond to the V_n for n odd. \square

REFERENCES

1. M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D.C., 1964.
2. C. A. Berenstein and L. Zalcman, Pompeiu's problem on symmetric spaces, *Comment. Math. Helvetici*, 55 (1980) 593–621.

3. T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Springer-Verlag, New York, 1985.
4. L. Brown and J. P. Kahane, A note on the Pompeiu problem for convex domains, *Math. Ann.*, 259 (1982) 107–110.
5. H. T. Laquer, A perturbation formula with applications to Pompeiu-type problems, preprint.
6. J. E. Marsden and A. J. Tromba, *Vector Calculus*, third Edition, W. H. Freeman and Company, New York, 1988.
7. R. S. Strichartz, Radon inversion—variations on a theme, *Amer. Math. Monthly*, 89 (1982) 377–384.
8. P. Ungar, Freak theorem about functions on a sphere, *J. Lond. Math. Soc.*, 29 (1954) 100–103.
9. S. A. Williams, A partial solution of the Pompeiu problem, *Math. Ann.*, 223 (1976) 183–190.
10. L. Zalcman, Offbeat integral geometry, *Amer. Math. Monthly*, 87 (1980) 161–175.
11. ———, Analyticity and the Pompeiu problem, *Arch. Rational Mech. Anal.*, 47 (1972) 237–254.

Department of Mathematics
Idaho State University
Pocatello, ID 83209
laquerht@csc.isu.edu

II. A DIGIT FOR NEGATIVE ONE.

BY J. P. BALLANTINE, COLUMBIA UNIVERSITY.

Mathematical historians will tell you how many years mathematics was held back for want of a digit 0. Though not comparing in importance with that digit, there is a certain advantage in having a digit to represent negative one. For this purpose we will use the digit for 1 inverted: thus I.

We have as a matter of fact no well-accepted way of writing a negative number, except by the ambiguous minus sign which usually denotes subtraction. It is no wonder that students do not grasp the logical difference between the problem in subtraction 0-7 and the number negative seven which we now may denote I3.

In numerical work with logarithms, one runs across such numbers as $9.69897 - 10$. How much simpler to write I9.69897 or I.69897? It may even be written I.70117.

The laws of operation of the new digit are easily mastered. In the new multiplication table, such entries as $I \times 7 = I3$ are easily memorized. Such identities as $I3 = I93 = I993 = I99993$ are also obvious. This latter remark is of significance in connection with computing machines. It is commonly understood that a big string of digits 9 extending to the left on the machine is a negative number, and is commonly explained by use of the words "complementary number." The whole thing becomes clear immediately if we place I at the head of these nines.¹

—*American Mathematical Monthly*
 32, (1925) p. 302.

A Quicker Convergence to Euler's Constant

Duane W. DeTemple

Euler's constant γ is usually defined by the limit relation

$$\gamma = \lim_{n \rightarrow \infty} D_n,$$

where

$$D_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n.$$

The rate of convergence is extremely slow, since

$$\frac{1}{2(n+1)} < D_n - \gamma < \frac{1}{2n}, \quad n = 1, 2, \dots$$

Young gave an elementary proof of this inequality in [2].

It is less well known that convergence can be improved significantly by replacing D_n with

$$R_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log\left(n + \frac{1}{2}\right).$$

Clearly $\lim_{n \rightarrow \infty} R_n = \gamma$; what is unexpected is the large effect this slight change has on the rate of convergence.

Theorem. For all natural numbers n ,

$$\frac{1}{24(n+1)^2} < R_n - \gamma < \frac{1}{24n^2}.$$

Proof: The upper and lower bounds are both obtained by considering the function f defined on $x > 0$ by

$$f(x) = -(x+1)^{-1} - \log\left(x + \frac{1}{2}\right) + \log\left(x + \frac{3}{2}\right).$$

Short calculations show that

$$R_n - R_{n+1} = f(n)$$

and

$$f'(x) = -\frac{1}{4}(x+1)^{-2}\left(x + \frac{1}{2}\right)^{-1}\left(x + \frac{3}{2}\right)^{-1}.$$

To obtain the upper bound, we first observe that

$$-f'(x) < \frac{1}{4}\left(x + \frac{1}{2}\right)^{-4}.$$

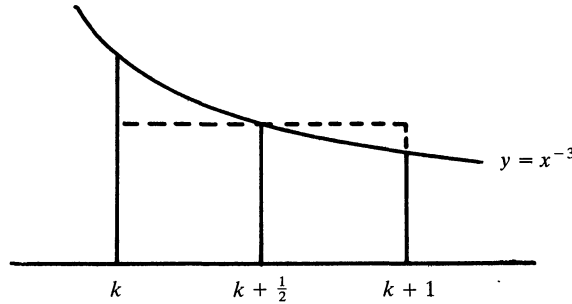
Therefore, since $f(\infty) = 0$,

$$f(k) = -\int_k^\infty f'(x) dx < \frac{1}{4} \int_k^\infty (x + \frac{1}{2})^{-4} dx = \frac{1}{12} (k + \frac{1}{2})^{-3}.$$

Since $(k + \frac{1}{2})^2 > k(k + 1)$, it follows that

$$\left(k + \frac{1}{2}\right)^{-3} < \frac{1}{2} \frac{2k + 1}{k^2(k + 1)^2} = \int_k^{k+1} x^{-3} dx.$$

Indeed it is geometrically evident from the figure that $(k + \frac{1}{2})^{-3}$ is smaller than the area beneath $y = x^{-3}$ over $k \leq x \leq k + 1$.



Altogether then,

$$\begin{aligned} R_n - \gamma &= \sum_{k=n}^{\infty} (R_k - R_{k+1}) = \sum_{k=n}^{\infty} f(k) \\ &< \frac{1}{12} \sum_{k=n}^{\infty} \left(k + \frac{1}{2}\right)^{-3} < \frac{1}{12} \int_n^{\infty} x^{-3} dx = \frac{1}{24n^2}. \end{aligned}$$

To derive the lower bound, we require the inequality

$$(x + \frac{1}{2})(x + \frac{3}{2}) = x^2 + 2x + \frac{3}{4} < (x + 1)^2,$$

from which it follows that

$$-f'(x) > \frac{1}{4}(x + 1)^{-4}.$$

Proceeding as before, we find that

$$f(k) > \frac{1}{4} \int_k^{\infty} (x + 1)^{-4} dx = \frac{1}{12} (k + 1)^{-3},$$

and then obtain

$$R_n - \gamma > \frac{1}{12} \sum_{k=n}^{\infty} (k + 1)^{-3} > \frac{1}{12} \int_{n+1}^{\infty} x^{-3} dx = \frac{1}{24(n + 1)^2}.$$

This completes the proof of the theorem. ■

The same ideas used in the proof above can be adapted to obtain the higher order estimate

$$R_n - \gamma - \frac{1}{24(n + \frac{1}{2})^2} = -r_n,$$

where

$$\left(\frac{7}{960}\right) \frac{1}{(n+1)^4} < r_n < \left(\frac{7}{960}\right) \frac{1}{n^4}.$$

The general expansion to arbitrary order can be found in terms of Bernoulli numbers. A derivation which requires only elementary calculus can be found in [1].

REFERENCES

1. D. W. DeTemple and S. H. Wang, Half integer approximations for the partial sums of the harmonic series, *J. Math. Analysis and Applic.* 160 (1991), 149–156.
2. R. M. Young, Euler's Constant, *Math. Gazette* 75, No. 472 (1991), 187–190.

*Department of Pure and Applied Mathematics
Washington State University
Pullman, Washington 99164-3113*



A Strange Attractor.

The Minimal Polynomial of $\cos(2\pi / n)$

William Watkins and Joel Zeitlin

It is an old result that the primitive n th root of unity, $\zeta_n = \cos(2\pi/n) + i \sin(2\pi/n)$, satisfies a rational polynomial of degree $\phi(n)$ —the number of integers between 1 and n that are relatively prime to n . The most common way to compute the minimal polynomial of ζ_n (called the n th cyclotomic polynomial and denoted by $\Phi_n(x)$) uses the identity [van, p. 114]

$$x^n - 1 = \prod_{d|n} \Phi_d(x). \quad (1)$$

Then $\Phi_9(x)$, for example, can be computed in the following familiar way:

$$x^9 - 1 = \Phi_1(x) \Phi_3(x) \Phi_9(x),$$

$$x^3 - 1 = \Phi_1(x) \Phi_3(x),$$

so that

$$\Phi_9(x) = \frac{x^9 - 1}{x^3 - 1} = x^6 + x^3 + 1.$$

Now consider the real and imaginary parts of ζ_n , i.e., $\cos(2\pi/n)$ and $\sin(2\pi/n)$. They too are algebraic numbers, but how can we compute *their* minimal polynomials? In 1933, D. H. Lehmer [Leh], [Niv] described a method for constructing these minimal polynomials from the cyclotomic polynomials. In this note, we present identities, analogous to (1), for the minimal polynomial $\Psi_n(x)$ of $\cos(2\pi/n)$, which provide another method for computing $\Psi_n(x)$. The identities involve the Chebychev polynomials $T_s(x)$, which are defined by

$$T_s(\cos \theta) = \cos(s\theta),$$

for positive integers s and all real θ . The degree of $T_s(x)$ is s and the leading coefficient is 2^{s-1} .

Theorem. *Let $\Psi_n(x)$ be the minimal polynomial of $\cos(2\pi/n)$ and let $T_s(x)$ denote the s th Chebychev polynomial.*

a) *If $n = 2s + 1$ is odd, then*

$$T_{s+1}(x) - T_s(x) = 2^s \prod_{d|n} \Psi_d(x), \quad (2)$$

and

b) *if $n = 2s$ is even, then*

$$T_{s+1}(x) - T_{s-1}(x) = 2^s \prod_{d|n} \Psi_d(x). \quad (3)$$

Before getting to the proof of the theorem, we show how (2) and (3) can be used to compute $\Psi_n(x)$ in the same way that identity (1) is used to compute $\Phi_n(x)$. For example, to compute $\Psi_9(x)$ (the minimal polynomial of $\cos(2\pi/9)$), let $s = 4$ and

then $s = 1$ in equation (2) to get

$$T_5(x) - T_4(x) = 16\Psi_1(x)\Psi_3(x)\Psi_9(x),$$

and

$$T_2(x) - T_1(x) = 2\Psi_1(x)\Psi_3(x).$$

Then

$$8\Psi_9(x) = \frac{T_5(x) - T_4(x)}{T_2(x) - T_1(x)}.$$

These four Chebychev polynomials are easy to compute. For example,

$$\begin{aligned} T_5(\cos \theta) &= \cos(5\theta) \\ &= \operatorname{Re}((\cos \theta + i \sin \theta)^5) \\ &= \cos^5 \theta - 10 \cos^3 \theta \sin^2 \theta + 5 \cos \theta \sin^4 \theta \\ &= \cos^5 \theta - 10 \cos^3 \theta (1 - \cos^2 \theta) + 5 \cos \theta (1 - \cos^2 \theta)^2 \\ &= 16 \cos^5 \theta - 20 \cos^3 \theta + 5 \cos \theta, \end{aligned}$$

so that

$$T_5(x) = 16x^5 - 20x^3 + 5x.$$

Similarly,

$$\begin{aligned} T_4(x) &= 8x^4 - 8x^2 + 1, \\ T_2(x) &= 2x^2 - 1, \end{aligned}$$

and

$$T_1(x) = x.$$

Thus

$$8\Psi_9(x) = \frac{T_5(x) - T_4(x)}{T_2(x) - T_1(x)} = 8x^3 - 6x + 1.$$

Incidentally, the theorem states that either $T_{s+1}(x) - T_s(x)$ or $T_{s+1}(x) - T_{s-1}(x)$ (depending on whether n is odd or even) is an annihilating rational polynomial for $\cos(2\pi/n)$. But a more obvious annihilating polynomial for $\cos(2\pi/n)$ is $T_n(x) - 1$, since $T_n(\cos(2\pi/n)) = \cos(2\pi) = 1$. These polynomials are related by the identities [Riv, p. 5]

$$\begin{aligned} (x - 1)(T_{2s+1}(x) - 1) &= (T_{s+1}(x) - T_s(x))^2, \\ 2(x^2 - 1)(T_{2s}(x) - 1) &= (T_{s+1}(x) - T_{s-1}(x))^2. \end{aligned}$$

Thus $\cos(2\pi/n)$ is a double root of $T_n(x) - 1$ whenever $n > 2$.

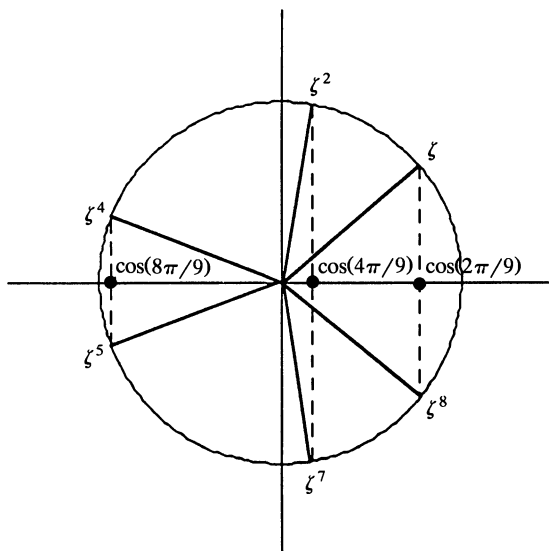
To begin the proof of the theorem, we need a few facts about the degree and the conjugates of $\cos(2\pi/n)$ over the rational numbers \mathcal{Q} . Since $\cos(2\pi/n) = (\zeta_n + \zeta_n^{-1})/2$, we have

$$\mathcal{Q}(\zeta_n) \supseteq \mathcal{Q}(\cos(2\pi/n)) \supseteq \mathcal{Q}.$$

To compute the degree of the extension $\mathcal{Q}(\cos(2\pi/n))$ over \mathcal{Q} , observe that $[\mathcal{Q}(\zeta_n) : \mathcal{Q}] = \phi(n)$ and $[\mathcal{Q}(\zeta_n) : \mathcal{Q}(\cos(2\pi/n))] = 1$ or 2 , since ζ_n is a root of the quadratic polynomial $x^2 - 2\cos(2\pi/n)x + 1$. Now if $n \geq 3$, ζ_n is not real and so $[\mathcal{Q}(\zeta_n) : \mathcal{Q}(\cos(2\pi/n))] = 2$. Thus we have:

$$\deg \Psi_n(x) = \begin{cases} 1, & \text{if } n = 1, 2 \\ \phi(n)/2, & \text{if } n \geq 3. \end{cases}$$

To make further progress, we need to find *all* the roots of $\Psi_n(x)$, i.e., the conjugates of $\cos(2\pi/n)$. For $n = 1, 2, 3, 4, 6$, $\cos(2\pi/n)$ is rational and so $\Psi_n(x)$ is linear and has just one root. But for other values of n , $\phi(n)/2 \geq 2$ and thus $\Psi_n(x)$ has more than one root. What are they? To take a specific example, what are the roots of the cubic polynomial $8\Psi_9(x) = 8x^3 - 6x + 1$? Of course, the roots of $\Phi_9(x)$ are $\zeta, \zeta^2, \zeta^4, \zeta^5, \zeta^7, \zeta^8$, where $\zeta = \zeta_9$.



They occur in pairs $\cos(2\pi k/9) \pm i \sin(2\pi k/9)$, $k = 1, 2, 4$, with only three distinct real parts: $\cos(2\pi/9)$, $\cos(4\pi/9)$, $\cos(8\pi/9)$. These three cosines are the roots of $\Psi_9(x)$. The general situation is this:

Lemma. *If $n \geq 3$, then the roots of $\Psi_n(x)$ are $\cos(2\pi k/n)$, for $0 \leq k \leq s$ and $(k, n) = 1$. (s is defined as in the theorem.)*

This lemma follows easily from the fact that the Q -automorphism σ of $Q(\zeta_n)$ given by $\sigma(\zeta_n) = \zeta_n^k$, for $(k, n) = 1$, sends $\cos(2\pi/n)$ to $\cos(2\pi k/n)$:

$$\sigma(\cos(2\pi/n)) = \sigma((\zeta_n + \zeta_n^{-1})/2) = (\zeta_n^k + \zeta_n^{-k})/2 = \cos(2\pi k/n).$$

Thus,

$$0 = \sigma(\Psi_n(\cos(2\pi/n))) = \Psi_n(\sigma(\cos(2\pi/n))) = \Psi_n(\cos(2\pi k/n))$$

and so the $\phi(n)/2$ numbers $\cos(2\pi k/n)$, where $(k, n) = 1$ and $0 \leq k \leq s$, are the roots of $\Psi_n(x)$.

Now to prove part a) of the theorem ($n = 2s + 1$ is odd), it suffices to show that both sides of (2) have the same roots and the same leading coefficient. The degree of the right side of (2) is

$$\sum_{d|n} \deg(\Psi_d(x)) = \deg(\Psi_1(x)) + \frac{1}{2} \sum_{d|n, d \neq 1} \phi(d) = 1 + \frac{1}{2}(n - 1) = s + 1,$$

which agrees with the degree of the left side of (2). The left side and the right side of (2) have exactly the same $s + 1$ roots: $\cos(2\pi k/n)$, $0 \leq k \leq s$, all of which are

distinct. To see this let $(k, n) = g$ so that $k = k'g$, $n = dg$, and $(k', d) = 1$. Then from the lemma, $\cos(2\pi k/n) = \cos(2\pi k'/d)$ is a root of $\Psi_d(x)$, which is a factor of the right side of (2). On the left side,

$$\begin{aligned} & T_{s+1}(\cos(2\pi k/n)) - T_s(\cos(2\pi k/n)) \\ &= \cos(2\pi k(s+1)/n) - \cos(2\pi ks/n) \\ &= \cos(\pi k(n+1)/n) - \cos(\pi k(n-1)/n) \\ &= \cos(\pi k + \pi k/n) - \cos(\pi k - \pi k/n) \\ &= 0. \end{aligned}$$

So $\cos(2\pi k/n)$ is also a root of the left side of (2). Now since the leading coefficient of $T_{s+1}(x)$ is 2^s , identity (2) is proved.

The proof of part b) is similar, but in this case the degree of the right side of (3) is computed as follows:

$$\begin{aligned} \sum_{d|n} \deg(\Psi_d(x)) &= \deg(\Psi_1(x)) + \deg(\Psi_2(x)) + \frac{1}{2} \sum_{d|n, d>2} \phi(d) \\ &= 2 + \frac{1}{2}(n-2) = s+1, \end{aligned}$$

which agrees with the degree of the left side of (3). Again the polynomials on the right and left sides of (3) have exactly the same $s+1$ roots: $\cos(2\pi k/n)$, $0 \leq k \leq s$, and the same leading coefficient. To see that $\cos(2\pi k/n)$ is a root of the left side of (3), compute

$$\begin{aligned} & T_{s+1}(\cos(2\pi k/n)) - T_{s-1}(\cos(2\pi k/n)) \\ &= \cos(2\pi k(s+1)/n) - \cos(2\pi k(s-1)/n) \\ &= \cos(\pi k(n+2)/n) - \cos(\pi k(n-2)/n) \\ &= \cos(\pi k + 2\pi k/n) - \cos(\pi k - 2\pi k/n) \\ &= 0. \end{aligned}$$

■

Note. K. W. Wegner [Weg] provides a list of the minimal polynomials (all of degree less than eight) for trigonometric functions of certain angles that are rational multiples of 2π . In his monograph, *Irrational Numbers* [Niv, p. 39], I. Niven gives a formula for the degree of the minimal polynomial of the tangent of any rational multiple of 2π .

REFERENCES

-
- [Leh] D. H. Lehmer, A note on trigonometric algebraic numbers, *American Mathematical Monthly*, 40 (1933) 165–166.
[Niv] Ivan Niven, *Irrational Numbers*, MAA, Washington, D.C., 1956.
[Riv] T. J. Rivlin, *Chebyshev Polynomials from Approximation Theory to Algebra and Number Theory*, 2nd ed., Wiley, New York, 1990.
[van] B. L. van der Waerden, *Modern Algebra* (English Translation), 2nd ed., Ungar, New York, 1953.
[Weg] Kenneth W. Wegner, Equations with trigonometric values as roots, *American Mathematical Monthly*, 66 (1959) 52–53.

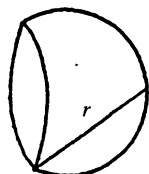
Department of Mathematics
California State University, Northridge
Northridge, CA 91330

The Equal Area Zones Property

B. Richmond and T. Richmond

A classical exercise found in many calculus texts is the *zones of a sphere* problem: Verify that if a sphere is sliced by two parallel planes h units apart, then the surface area of the zone between the planes is dependent on h alone, and independent of the location of the planes. While this property of the sphere is usually surprising to students, it has long been known. It is an immediate consequence of the following result of Archimedes (see p. 185ff in [1]) illustrated by the figure below.

Proposition. *The surface area of any segment of a sphere is equal to the area of a circle whose radius is equal to the line drawn from the vertex of the segment to a point on the base circle of the segment of the sphere.*



$$\text{Surface Area} = \pi r^2$$

Viewing the sphere as a surface of revolution with axis of symmetry perpendicular to the parallel planes slicing it, it is natural to extend the routine exercise mentioned above by asking the question of which other surfaces of revolution enjoy this equal area zones property. Clearly the cylinder does. Are there others? Approaching this problem from different viewpoints, we present two solutions which are easily accessible to undergraduates.

To state the problem precisely, suppose $y = g(x)$ is a piecewise smooth nonnegative curve defined over $[a, b]$, and is revolved around the x -axis. A *zone of width* h ($h \leq b - a$) of the resulting surface is the portion of the surface bounded between planes $x = x_0$ and $x = x_0 + h$, where $x_0 \in [a, b - h]$. Which surfaces of revolution have the *equal area zones* (E.A.Z.) *property* that for any width $h \in [0, b - a]$, the surface area

$$S(x, h) = \int_x^{x+h} 2\pi g(t) \sqrt{1 + g'(t)^2} dt \quad (x \in [a, b - h])$$

of a zone of width h is a function of h alone, independent of x ?

Two observations might be made about the two surfaces we know to have the E.A.Z. property, the sphere and the cylinder. First, one might observe that their generating curves, the circle and the line, are the only planar curves with nonnegative constant curvature. Secondly, in verifying that these two surfaces have the

E.A.Z. property, one discovers that in both cases, the integrand $f(t) = 2\pi g(t)\sqrt{1 + g'(t)^2}$ is constant. This is clearly a sufficient condition for the E.A.Z. property. In our first solution, we will show that, in the case of a smooth curve $g(t)$, it is also necessary.

Suppose $g(t)$ is smooth and nonnegative over $[a, b]$. For a fixed h , the surface area integral $S(x, h)$ above has the form $S(x) = \int_x^{x+h} f(t) dt$ where $f(t) = 2\pi g(t)\sqrt{1 + g'(t)^2}$. If the surface generated by $g(t)$ has the E.A.Z. property, then for a fixed h , $S(x)$ is constant, so

$$\frac{d}{dx}S(x) = \frac{d}{dx} \int_x^c f(t) dt + \frac{d}{dx} \int_c^{x+h} f(t) dt = 0.$$

Applying the Fundamental Theorem of Calculus, we find that $f(x) = f(x + h)$ for any $x \in [a, b - h]$, so $f(x)$ is h -periodic over $[a, b]$. Since this argument holds for any $h \in [0, b - a]$, it follows that the integrand $f(t)$ must be constant.

Now to determine for which curves $y = g(x)$ the integrand $f(x)$ is constant, we must solve the differential equation

$$y\sqrt{1 + y'^2} = c.$$

If $c = 0$, then $g(x) \equiv 0$, so we may assume $c \neq 0$. It follows that $y = g(x) \neq 0$ for any x , and upon squaring both sides of the equation and solving for y' , we get

$$y' = \frac{dy}{dx} = \pm \sqrt{\frac{c^2}{y^2} - 1}. \quad (1)$$

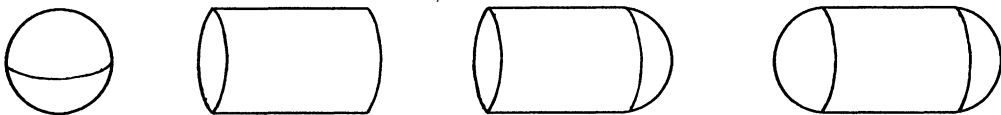
If $y \equiv c$ in some interval, we have $y = g(x)$ is constant, and the corresponding surface over that interval is a cylinder. If $g(x_0) \neq c$ for some $x_0 \in [a, b]$, by continuity, we have $g(x) \neq c$ in some interval I containing x_0 , and separation of variables gives

$$\left[\left(\frac{c}{y} \right)^2 - 1 \right]^{-1/2} dy = \pm dx.$$

Substituting $u = c/y$ and integrating gives

$$\int \frac{c du}{u^2 \sqrt{u^2 - 1}} = \pm \int dx.$$

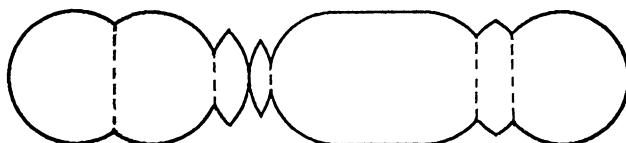
The trigonometric substitution $u = \csc \theta$ yields $cu^{-1}\sqrt{u^2 - 1} = \pm x + k$, or $y^2 + (\pm x + k)^2 = c^2$. From this form, it is clear that the corresponding curve over I is a portion of a semicircle of radius c centered at $(\pm k, 0)$. This, together with the smoothness of g implies that if $g(x_1) = c = g(x_2)$ for $x_1 < x_2$, then $g(x) = c$ for all $x \in [x_1, x_2]$. We conclude that revolving a nonnegative smooth curve yields a surface with the E.A.Z. property iff the surface is a sphere, a cylinder, a "silo", or a "capsule."



Only Smooth Surfaces with E.A.Z. Property

These four surfaces can all be described as zones of a capsule, where a capsule is a cylinder of height $l \geq 0$ with a hemisphere attached to each end. Alternately, the differential equation (1) could be solved by the substitution $u = y^2$, which leads to a separable differential equation which can be integrated using only the power rule. The solutions of (1) above also illustrate the fact that the envelope of a family of solutions to a first order differential equation is again a solution to the differential equation. (e.g., see [2].) The envelope of the spheres of radius c centered at $(\pm k, 0)$ is the cylinder of radius c .

It follows that if $g(x)$ is a continuous nonnegative piecewise smooth curve that generates a surface of revolution with the E.A.Z. property, then the surface must be a "string of beads" where every bead is a zone of a capsule of fixed radius c .



A String of Zones of a Capsule

Rather than a direct proof as presented above, with enough intuition one might pursue a uniqueness argument to show that the zones of a capsule are the only smooth surfaces with the E.A.Z. property. Such an argument follows.

Suppose g is a smooth nonnegative curve over $[a, b]$ generating a surface of revolution with the E.A.Z. property, and with the same zone surface area function $S(h) = 2\pi rh$ as the sphere generated by $c(x) = \sqrt{r^2 - x^2}$. If $g(x_0) > r$, then $2\pi g(x_0)\sqrt{1 + g'(x_0)^2} > 2\pi r = 2\pi c(x_0)\sqrt{1 + c'(x_0)^2}$, and by continuity, there exists $\varepsilon > 0$ such that $2\pi g(x)\sqrt{1 + g'(x)^2} > 2\pi c(x)\sqrt{1 + c'(x)^2}$ for all $x \in I = [x_0 - \varepsilon, x_0 + \varepsilon] \cap [a, b]$. This gives the contradiction that the surface area of any zone generated by g within the interval I has greater area than a zone of equal width generated by $c(x)$. Thus, $g(x) \leq r$ for all x , and therefore, for any x_0 , there exists z_0 such that $g(x_0) = c(z_0)$ and $g'(x_0)$ and $c'(z_0)$ do not have opposite signs. If $g'(x_0) \neq c'(z_0)$, then as above we conclude that there exists $\varepsilon > 0$ such that

$$\left| 2\pi g(x_0 + \delta)\sqrt{1 + g'(x_0 + \delta)^2} - 2\pi c(z_0 + \delta)\sqrt{1 + c'(z_0 + \delta)^2} \right| > 0$$

for $|\delta| < \varepsilon$,

contradicting that g and c generate surfaces with zones of equal area. Thus $g'(x_0) = c'(z_0)$ whenever $g(x_0) = c(z_0)$. It follows that on any interval on which g is invertible, that is, on which $g'(x) \neq 0$, or equivalently, $g(x) \neq r$, g agrees with a horizontal translation of a portion of the curve c . Using a smoothness argument as in the previous solution, we conclude that the surface generated by g must be a zone of a capsule.

REFERENCES

1. E. J. Dijksterhuis, *Archimedes*, Princeton University Press, Princeton, N.J., 1987.
2. Ray Redheffer, *Differential Equations, Theory and Applications*, Jones and Bartlett Publishers, Boston, 1991.

Department of Mathematics
Western Kentucky University
Bowling Green, KY 42101

Graph Theory and the Game of Sprouts

Mark Copper

This article concerns a game that children can play. As described by Martin Gardner [1] it was invented and studied by John H. Conway and Michael S. Paterson. This account has written it as a digression on Euler's formula that one might present to an undergraduate graph theory class.

Sprouts is a game played with pen and paper. In it the players begin with a finite set of m points. For his turn a player adds both an edge and a vertex to the graph in a prescribed way. Namely, the player must first add an edge between two extant vertices (or add a loop at a single vertex) and then subdivide the edge he just added with a new vertex. To make the game finite, we require that no vertex ever be of degree greater than 3, and to make it interesting, we prohibit edges from crossing one another. The winner of the game is, say, the player who makes the last possible move.

In what follows we consider upper and lower bounds on the number of plays in a complete game of Sprouts on m points. In the first proposition, it is shown that a game may last $3m - 1$ plays but no more. In Propositions 2 and 3 it is shown that a game may last $2m$ plays but no less. It is shown in Proposition 5 that a game whose final graph is connected requires strictly more than $2m$ plays while in Proposition 4 it is shown that if the final graph is 2-connected, the number of plays required is not less than $(7m - 2)/3$.

There is only one game possible if $m = 1$. Although such games may be drawn in apparently different ways (see FIGURE 1), each can be identified with a game

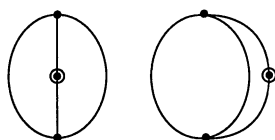


Figure 1

played on a sphere, this by identifying the plane (homeomorphically) with a punctured sphere. Once on the sphere any game on one point can be deformed to any other. There are two different games on two points shown in FIGURE 2. Note that the game in (2a) has stopped because there is no legal way to connect the two vertices of degree 2 (the circled vertices) without crossing an edge. As we will see,

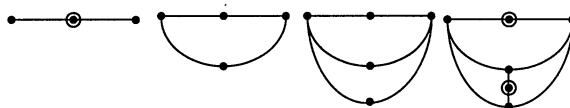


Figure 2a

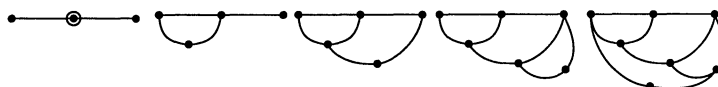


Figure 2b

these two games are extremal in the sense that every game on two points must last at least four moves and no more than five.

Proposition 1. *A game on m points can last no longer than $3m - 1$ plays, and a game of that length is always possible.*

Proof: Suppose a game starts with m points. After p plays there are $m + p$ vertices, and since each play adds four to the total degree of the graph, the total degree after p plays is $4p$. According to the rules of play, no vertex is allowed to have degree larger than 3. Hence after p plays the total degree will not exceed $3(m + p)$. Actually, it must be strictly less since there is always a vertex of degree 2 after play has started. Thus $4p < 3(m + p)$ or $p < 3m$.

We will now describe a game that takes $3m - 1$ plays to finish. First note that when a game on one point is played, the plane is divided into 3 regions, and the vertex of degree 2 lies on the boundary of precisely two of these regions. Now fix a point and play a game on that point so that after the two plays all the remaining points are contained in one of the regions with the degree 2 vertex in its boundary. This is the first step. For the second, play the one point game at a second vertex in such a way that all the remaining degree zero vertices as well as the first one-game lie in a single region with the degree 2 vertex of the second one-game in its boundary. With the next play, connect the two vertices of degree 2. See FIGURE 3. The “dumbbell” so constructed contains a single vertex of degree 2 and it lies in the boundary of the region containing the remaining null-graph. That is, we are effectively in the same situation as after the first step, and we can repeat the second step until all the points are used up. The first step took two plays and the second three. Hence the total will be $3m - 1$. This proves the proposition.

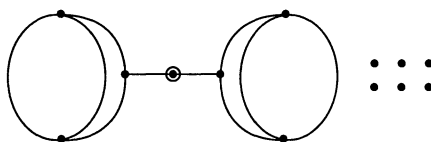


Figure 3

Proposition 2. *There is a complete game of Sprouts on m points which takes exactly $2m$ plays.*

Proof: As in the long game described in Proposition 1, choose a point and play a one-game on it. Keep the null-points in a single region, but this time put them in the region that does *not* contain the degree 2 vertex in its boundary. This isolates the degree 2 vertex, and it must remain degree 2 for the remainder of the game.

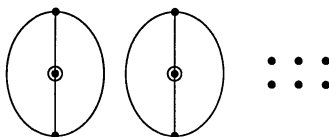


Figure 4

Repeat this step for each of the remaining points. See FIGURE 4. This process yields a game that terminates in $2m$ plays.

As we will explain, not only is this a game of minimal length, but every other game must last longer. First, however, let us recall Euler's formula for planar graphs. Let G be any graph; suppose it has e edges and v vertices. We say that G is planar if it can be drawn in the plane so that none of its edges cross. Since the 2-plane is homeomorphic to the 2-sphere with a single point removed, we may also think of a planar graph as one arising from some polyhedron in 3-space. The edges of a planar graph divide the plane into regions which, thinking of the associated polyhedron, we will call faces. Let f be the number of faces in G . Euler's formula relates these quantities:

$$f - e + v = 2.$$

The figures that arise from a game of sprouts are not always graphs in the standard sense of the word since there may be multiple edges between vertices as in FIGURE 1. In the next paragraph, we will modify these figures so that even loops occur. Nevertheless Euler's formula remains valid in each instance, and we shall continue to call the figures under consideration "graphs."

Suppose that a graph G_0 has been obtained from a complete game of sprouts. Each vertex of G_0 is of degree 2 or 3. This graph is a "subdivision" of a unique cubic graph G which we may obtain from G_0 as follows. Suppose x is a vertex of degree 2 and suppose that x is adjacent to y . Remove an edge connecting x and y , then identify the vertices x and y . Do this once at each degree 2 vertex. That is, we contract one edge incident to each degree 2 vertex. Although we obtain a cubic graph (i.e., each vertex is of degree 3), we don't want to forget where the degree 2 vertices were, so we color each of the edges which before had been incident to a degree 2 vertex; red, say. In the figures, a degree 2 vertex is to be understood as an indication that the two incident edges are to be considered as a single red edge. Note that this process of contraction can produce loops in G , but that such loops will always be colored red.

Lemma 1. *Suppose that the cubic graph G arises as just described from a complete game of Sprouts played on m vertices in p plays. Then*

$$f = 2 + p - m.$$

Proof: Let r be the number of red edges in G . Reasoning as in Proposition 1, the total degree of G_0 is $4p$, and it is also $3(m + p) - r$. In particular, $r = 3m - p$. The number of edges in G , e , is $2p - r = 3(p - m)$. The number of vertices, v , is $m + p - r = 2(p - m)$. The lemma follows by substituting these values for e and v into Euler's formula for G .

Recall that a graph is connected if there is a path between any pair of vertices, and that an edge in a connected graph is called a bridge if its removal disconnects

the graph. The next proposition shows that the game in Proposition 2 is of minimal length.

Proposition 3. *Suppose that the cubic graph G arises from a complete game of Sprouts on m vertices in p plays. Then*

$$p \geq 2m.$$

Proof: We may assume that G is connected since the general case follows easily from this. Observe that no face in the graph G can have more than one red edge, for otherwise the generating game of Sprouts would not have been complete. Hence $f \geq r$. Since $r = 3m - p$, it follows from Lemma 1 that $2 + p - m \geq 3m - p$, and hence that

$$p \geq 2m - 1.$$

To finish the proof we must show that this inequality is strict. Suppose therefore that $p = 2m - 1$. Then, again by Lemma 1, $f = m + 1$. On the other hand, $r = m + 1$ as well. Thus each face must have a red edge and each red edge must lie in the boundary of a single face. Now an edge lies in the boundary of one or two edges according to whether or not it is a bridge. Thus each face in G has a bridge in its boundary. But this is impossible: G is not a tree since it is cubic and all trees have end vertices (vertices of degree 1). Thus G has at least one cycle. Since G is finite there must be a cycle with no other cycle in its interior. The interior of this cycle is a face of G which can have no bridge in its border without also having an end vertex. This proves the proposition.

What is the shortest game on m points whose final graph is connected? I don't know. The next two propositions give some information in this direction, however. In the first we glean a little more information from Euler's formula, and in the second we establish our claim that any game of $2m$ plays on m points must be the game described in Proposition 2. We say that a graph is 2-connected if it is connected and it contains no bridges.

Proposition 4. *Suppose that the cubic graph G arises from a complete game of Sprouts on m vertices in p plays. If G is 2-connected, then*

$$p \geq \frac{7}{3}m - \frac{2}{3}. \quad (*)$$

Proof: Since each edge lies in the boundary of exactly two faces, there must be twice as many faces as red edges. Thus, since $f \geq 2r$, we obtain

$$2 + p - m \geq 2(3m - p),$$

which simplifies to (*).

It remains for us to consider graphs which arise from a game of Sprouts which are connected but not necessarily 2-connected.

Proposition 5. *Suppose that the cubic graph G arises from a complete game of Sprouts on m vertices in p plays. If G is connected and $m > 2$, then*

$$p > 2m.$$

Proof: In light of Proposition 3 we need only show $p \neq 2m$. Suppose then that $p = 2m$. Then $f = 2 + m$ by Lemma 1. Let b be the number of red bridges in G . We have $m \geq b \geq m - 2$. If all the red bridges are removed from G , the connected components include at least two nontrivial subgraphs to which only one

bridge was attached. Such subgraphs cannot be loops since loops must be red, and no two red edges can be adjacent. Hence we can contract by an edge incident to the vertex where the bridge was attached and obtain a cubic planar graph. In such graphs $f \geq 3$ since $2e = 3v$. In particular, there are at least two interior faces. Thus

$$f = m + 2 \geq b + 4,$$

and b is actually equal to $m - 2$. Moreover, there must be exactly two end components when the red bridges are deleted and each of these must have exactly two interior faces. In the original graph G such a component has four edges, exactly one of which is red. But if we recall how the game is played, we realize that there must be an even number of edges in any such component which are not colored. Consequently, this configuration is also impossible, and the proposition follows.

Given the nature of the subject it should be no surprise that more questions have been raised than settled. It would certainly be more satisfying to have sharp lower bounds on the number of plays in a connected or a 2-connected game. It would also be interesting to know what happens to such bounds if Sprouts is played on some other manifold since we have relied so heavily on the Euler characteristic ($f - e + v$) of the sphere. One might also wonder which cubic planar graphs G can arise from a game of Sprouts. That is, when can G be decomposed into an edge sum of P_1 and P_2 subgraphs in such a way that no two of the P_1 summands bound the same face? (P_1 and P_2 denote the path graphs of one and two edges respectively.) Finally, since G is planar, it has a dual graph G^* in which the roles of vertices and faces are interchanged. In the context of G^* , we want to know when a set of edges is maximal with respect to both independence and the rules of the game. If G is 2-connected, for example, Tutte's 1-factor theorem applies to G^* , and gives a condition under which the inequality of Proposition 4 is strict.

REFERENCE

1. M. Gardner, *Mathematical Carnival*, Alfred A. Knopf, New York, 1975.

Department of Mathematics
Florida International University
Miami, FL 33199
copper@servax.fiu.edu

"The sine curve of Bush's presidency was nearly as predictable as geometry, if his Campaign behaviour is taken as the axiom."

from an article "All the President's Wars"
 by Sidney Blumenthal.

New Yorker, Dec. 28, 1992-Jan. 4, 1993, p. 66 lines 8-12.

—Contributed by Emma Lehmer

NOTES

Edited by: John Duncan

Elementary Proof of the Remez Inequality

Borislav Bojanov

This note is concerned with the Tchebycheff polynomials $T_n(x)$. As well known they can be presented on $[-1, 1]$ by the expression

$$T_n(x) = \cos(n \arccos x).$$

The famous Russian mathematician Pafnutii Lvovich Tchebycheff (1821–1894) introduced $T_n(x)$ as the polynomial of least uniform norm on $[-1, 1]$ amid the polynomials of degree n with fixed leading coefficient.

The Tchebycheff polynomials appear prominently in various extremal problems posed in π_n (the set of all polynomials of degree n). An illuminating example is the classical Markov inequality, which shows that

$$\|p^{(k)}\| \leq \|T_n^{(k)}\|, \quad k = 0, \dots, n,$$

for each $p \in \pi_n$ such that

$$\|p\| := \max\{|p(x)| : x \in [-1, 1]\} \leq 1.$$

The proof of this and many other remarkable properties of T_n can be found in the recent book of Rivlin [4].

It has been mentioned already by Tchebycheff that T_n is the fastest growing polynomial outside $[-1, 1]$. In other words,

$$\max\{|p(\xi)| : p \in \pi_n, \|p\| \leq 1\} = T_n(\xi)$$

for each $|\xi| \geq 1$. This observation provokes the following question: How large can a polynomial be given that it is constrained to be “small” on a substantial portion of its domain? Make the problem more precise as follows.

Let σ be an arbitrary fixed positive number. For every $p \in \pi_n$ define the set

$$M(p) := \{x \in [-1, 1 + \sigma] : |p(x)| \leq 1\}.$$

Clearly $M(p)$ consists of mutually disjoint closed subintervals. Let $|M(p)|$ be the measure of $M(p)$, i.e., $|M(p)|$ is the total length of these subintervals. Denote

$$\pi_n(\sigma) := \{p \in \pi_n : |M(p)| \geq 2\}.$$

The problem is to characterize the polynomial p^* from $\pi_n(\sigma)$ which has a maximal uniform norm over $[-1, 1 + \sigma]$.

Evidently, the Tchebycheff polynomial $T_n(x)$ belongs to $\pi_n(\sigma)$ for each $\sigma > 0$ since $|T_n(x)| \leq 1$ on $[-1, 1]$ and $|T_n(x)| > 1$ for $|x| > 1$. In 1936 Remez [1]

established the following

$$\sup_{p \in \pi_n(\sigma)} \|p\|_\infty = \|T_n\|_\infty, \quad (1)$$

where the supremum norm is over $[-1, 1 + \sigma]$. Of course $\|T_n\|_\infty = T_n(1 + \sigma)$. The proof of (1) can be seen also in the book of Freud [2]. A simpler approach was found recently by Erdelyi [3]. We demonstrate here a short, elementary proof.

The proof: Note that for any fixed $x \in [-1, 1 + \sigma]$ the quantity

$$\mu(x) := \sup\{|p(x)| : p \in \pi_n(\sigma)\}$$

is attained for some polynomial from $\pi_n(\sigma)$. We shall show first that $\mu(x) \leq \mu(1 + \sigma)$ for each $x \in [-1, 1 + \sigma]$. Indeed, let x be an interior point of $[-1, 1 + \sigma]$ and let p be the extremal polynomial for this point, i.e., $p \in \pi_n(\sigma)$ and $|p(x)| = \mu(x)$. Introduce the polynomials

$$p_1(x) := p(\alpha(x)), \quad p_2(x) := p(\beta(x)),$$

where $\alpha: [-1, 1 + \sigma] \rightarrow [-1, x]$ and $\beta: [-1, 1 + \sigma] \rightarrow [x, 1 + \sigma]$ are the linear transformations. Let M_1 and M_2 be the parts of $M(p)$ situated in $I_1 := [-1, x]$ and $I_2 := [x, 1 + \sigma]$, respectively. Assuming that $|M_i| < \lambda |I_i|$ for $i = 1, 2$ and $\lambda = 2/(2 + \sigma)$ we would get $|M| = |M_1 + M_2| < \lambda |I_1 + I_2| = \lambda(2 + \sigma) = 2$, a contradiction. Therefore $|M_i|/|I_i| \geq \lambda$ at least for one i , say for $i = 1$. Then $|M(p_1)| \geq 2$ and hence $p_1 \in \pi_n(\sigma)$. This yields

$$\mu(x) = |p(x)| = |p_1(1 + \sigma)| \leq \mu(1 + \sigma).$$

Therefore the Remez inequality will be proved if we show that

$$|p(1 + \sigma)| \leq T_n(1 + \sigma) \quad \text{for each } p \in \pi_n(\sigma).$$

In order to show this, denote by $-1 = \eta_0 < \eta_1 < \dots < \eta_n = 1$ the extremal points of T_n . We have

$$T_n(\eta_k) = (-1)^{n-k} \quad k = 0, \dots, n. \quad (2)$$

Let $x_0 < x_1 < \dots < x_n$ be the points of $M(p)$ which coincide with η_0, \dots, η_n after we press $M(p)$ to the left, i.e., to the interval $[-1, M(p) - 1]$. By the Lagrange interpolation formula

$$|p(1 + \sigma)| \leq \sum_{k=0}^n \prod_{\substack{i=0 \\ i \neq k}}^n \frac{|1 + \sigma - x_i|}{|x_k - x_i|}$$

since $|p(x_i)| \leq 1$. Now taking into account the obvious inequalities $|1 + \sigma - x_i| \leq |1 + \sigma - \eta_i|$, $|x_k - x_i| \geq |\eta_k - \eta_i|$ and (2), we get

$$|p(1 + \sigma)| \leq \sum_{k=0}^n \prod_{\substack{i=0 \\ i \neq k}}^n \frac{|1 + \sigma - \eta_i|}{|\eta_k - \eta_i|} = T_n(1 + \sigma).$$

The proof is completed.

The author is grateful to the referee and to the editor for their useful remarks.

1. E. J. Remez, Sur une propriété de polynomes de Tchebysheff, *Communications le l'Inst. des Sci.*, Kharkov 13 (1936), 93–95.
2. G. Freud, *Orthogonal Polynomials*, Pergamon Press, Oxford, 1971.
3. T. Erdelyi, *Inequalities for generalized polynomials and their applications*, Ph.D. Thesis, University of South Carolina, 1989.
4. T. J. Rivlin, *Chebyshev Polynomials*, second edition, John Wiley & Sons, Inc., New York, 1990.

*Department of Mathematics,
University of Sofia,
Boul. James Boucher 5,
1126 Sofia, BULGARIA*

A Note on an Identity of Ramanujan

T. S. Nanjundiah

In a forthcoming paper [1], Berndt and Bhargava have supplied a proof of this eye-catching identity of Ramanujan found in his third notebook [3, p. 386]: if $ad = bc$, then

$$\begin{aligned}
 & 64\{(b+c+d)^6 - (a+c+d)^6 - (a+b+d)^6 + (a+b+c)^6 \\
 & \qquad \qquad \qquad + (a-d)^6 - (b-c)^6\} \\
 & \times \{(b+c+d)^{10} - (a+c+d)^{10} - (a+b+d)^{10} \\
 & \qquad \qquad \qquad + (a+b+c)^{10} + (a-d)^{10} - (b-c)^{10}\} \\
 & = 45\{(b+c+d)^8 - (a+c+d)^8 - (a+b+d)^8 \\
 & \qquad \qquad \qquad + (a+b+c)^8 + (a-d)^8 - (b-c)^8\}^2.
 \end{aligned}$$

It figures also in their expository article [2] featuring a selected group of Ramanujan's results. Unfortunately, they have missed its simple proof and so its genesis by not noticing that it is built from two sets of sums:

$$\begin{aligned}
 u_n &= \alpha_1^n + \beta_1^n + \gamma_1^n, & \alpha_1 &= b+c+d, & \beta_1 &= -(a+b+c), & \gamma_1 &= a-d, \\
 v_n &= \alpha_2^n - \beta_2^n + \gamma_2^n, & \alpha_2 &= a+c+d, & \beta_2 &= -(a+b+d), & \gamma_2 &= b-c.
 \end{aligned}$$

By $\alpha_j + \beta_j + \gamma_j = 0$, the underlying problem is to compute

$$\omega_n = \alpha^n + \beta^n + \gamma^n,$$

where α , β and γ are the roots of the cubic

$$z^3 - pz + q = 0.$$

It is simple to work out an easy special case of Newton's formulae for power sums of the roots of an algebraic equation. Indeed, the obvious recursion

$$\omega_{n+3} - p\omega_{n+1} + q\omega_n = 0$$

with the initial values

$$\omega_{-1} = \frac{p}{q}, \quad \omega_0 = 3, \quad \omega_1 = 0,$$

yields

$$\begin{aligned} \omega_2 &= 2p, & \omega_4 &= 2p^2, \\ \omega_3 &= -3q, & \omega_5 &= 5pq, & \omega_7 &= -7p^2q, \\ \omega_6 &= 2p^3 + 3q^2, & \omega_8 &= 2p^4 + 8pq^2, & \omega_{10} &= 2p^5 + 15p^2q^2. \end{aligned}$$

Form the cubic whose roots are α_j , β_j and γ_j :

$$z^3 - p_j z + q_j = 0.$$

We have

$$\begin{aligned} p_1 &= (b + c + d)(a + b + c) + (a - d)^2, \\ p_2 &= (a + c + d)(a + b + d) + (b - c)^2, \\ p_1 - p_2 &= 3(bc - ad). \end{aligned}$$

Hence $p_1 = p_2$ if and only if

$$ad = bc.$$

Assume this condition and set

$$p_1 = p_2 = P, \quad \Delta = q_1^2 - q_2^2.$$

Now the $u_n = \omega_n(p_1, q_1)$ and the $v_n = \omega_n(p_2, q_2)$ given by the computed $\omega_n = \omega_n(p, q)$ show that

$$\begin{aligned} u_2 &= v_2, & u_4 &= v_4, \\ u_6 - v_6 &= -3\Delta, & u_8 - v_8 &= 8P\Delta, & u_{10} - v_{10} &= 15P^2\Delta. \end{aligned}$$

So we have Ramanujan's ingenious parametric construction of equal sums of three n th powers ($n = 2, 4$), and Ramanujan's identity. Clearly, for both these results, the condition $ad = bc$ is crucial. Ramanujan must have been primarily looking for the first one because of its number-theoretic significance, the second being incidental and apparently the only one of its kind in this context.

For special choices of the parameters, the equal sums of three n th powers ($n = 2, 4$) constructed by Ramanujan may present the same terms! This happens, for instance, when

$$a = b (c = d), \quad a = c (b = d), \quad b = 0 = d (a \neq 0), \quad c = 0 = d (a \neq 0).$$

Barring such cases, the construction yields numbers expressible as sums of three n th powers ($n = 2, 4$) in two different ways. This observation, which we owe to a comment of the referee/editor, does not point to any flaw in the construction for which what really matters is its *algebraic* formulation.

I wish to thank Professor Bhargava for having kindly shown me the proof sheets of [1] and a preprint of [2].

REFERENCES

1. Bruce C. Berndt and S. Bhargava, A remarkable identity found in Ramanujan's third notebook, *Glasgow Math. J.* 34 (1992) 341–345.
2. Bruce C. Berndt and S. Bhargava, Ramanujan—for Lowbrows, this MONTHLY (to appear).
3. S. Ramanujan, *Notebooks* (2 vols.), Tata Institute of Fundamental Research, Bombay, 1957.

180, I Cross
Gangotri Layout–I Stage
Mysore-570 009
India

On an Identity of Daubechies

Doron Zeilberger

Tossing a coin (whose $\Pr(\text{head}) = p$) until reaching n heads or n tails and equating the probability, 1, of finishing with the sum of the probabilities of all the possible final outcomes leads to

$$\sum_{i=0}^{n-1} \binom{n+i-1}{i} p^n (1-p)^i + \sum_{i=0}^{n-1} \binom{n+i-1}{i} p^i (1-p)^n = 1,$$

which was proved in [1], (pp. 167–171) and [2] using Bezout's theorem and induction respectively. Rolling a k -faced die instead leads to the multivariate generalization

$$\sum_{i=1}^k \sum_{\substack{0 \leq \alpha_j \leq n-1 \\ j \neq i}} \frac{(\alpha_1 + \cdots + \alpha_{i-1} + (n-1) + \alpha_{i+1} + \cdots + \alpha_k)!}{\alpha_1! \cdots \alpha_{i-1}! (n-1)! \alpha_{i+1}! \cdots \alpha_k!} \times$$

$$p_1^{\alpha_1} \cdots p_{i-1}^{\alpha_{i-1}} p_i^n p_{i+1}^{\alpha_{i+1}} \cdots p_k^{\alpha_k} = 1,$$

provided $p_1 + \cdots + p_k = 1$.

REFERENCES

1. Ingrid Daubechies, Ten lectures on wavelets, *SIAM*, Philadelphia, 1992.
2. ———, Orthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, 41 (1988), 909–996.

Department of Mathematics
Temple University
Philadelphia, PA 19122
zeilberg@euclid.math.temple.edu

COMPUTER SCIENCE SAMPLER

Edited by: Catherine C. McGeoch

In the May 1968 issue of the Monthly, G. E. Forsythe wrote an article titled "What to do till the computer scientist comes." Among his several suggestions: "Read [about computer science]. Since computer science is not yet very deep and mathematicians are very smart people, this should not be onerous."

There is no doubt that the ties between mathematics and computer science are strong and that each field has had a profound influence upon the other. It is also true that computer science (and mathematics) has changed and developed considerably since the 1960's. In the "Computer Science Sampler" I and guest columnists will try and give a glimpse of what computer scientists have been up to lately: we will write about intriguing mathematical results, old and new, that make possible the development of modern computing machines and computational methods.

Although computer science has gotten a lot "deeper" in the 20 years since Forsythe's article, reading the columns should not be onerous. After all, mathematicians are still very smart people.

Data Compression

Catherine C. McGeoch

Every object stored in a computer, whether an integer, the text of the *Oxford English Dictionary*, or a digitized image of the Mona Lisa, must first be *encoded* into a sequence of 0's and 1's (called bits). Alphabetic characters are usually represented according to either the ASCII (ask-ee) or the EBCDIC (ib-se-dic) standard code. For example, "A" is encoded 01000001 in ASCII and 11000001 in EBCDIC.

Suppose you want to store the text of *Far from the Madding Crowd* by Thomas Hardy. The book contains 768,771 characters: since both standard codes use 8 bits (one byte) per character the book would occupy slightly over half of a 3.5 inch floppy disk. Methods of *data compression* can be applied so that Hardy's book requires an average of 2.48 bits per character [1], thereby reducing the storage requirements by a factor of three.

Samuel Morse used a form of data compression in the design of his famous code. The frequently used letters have short sequences (E and T are · and - ·), and the less common letters have long sequences (Y and Z are - · - · and - - · ·). Although an alphabet of 30 characters requires $3.26 = [2 \cdot 1 + 4 \cdot 2 + 8 \cdot 3 + 14 \cdot 4]/30$ bits per character on average (using two 1-bit codes, four 2-bit codes, and so on), we might expect that a message in Morse Code would be shorter than

average because letters with short codes appear frequently. In this column we shall examine a data compression scheme that produces optimally-short encodings.

First, some definitions. An *alphabet* is a finite set of characters. We will denote a special alphabet $\beta = \{0, 1\}$. A *word* is a finite sequence of characters from some alphabet. (Although examples in this column will use “natural English” words, this need not be the case in general.) A *message* is a sequence of words. A *code* C is a one-to-one onto function mapping a set of *source words* $W = \{w_1, w_2, \dots, w_n\}$ from some alphabet to a set of *code words* $\{b_1, \dots, b_n\}$ from β . We *encode* a sequence of source words by applying C to each word in sequence. We *decode* a message by applying the inverse function C' to the coded words.

A *prefix code* is one in which no code word is a proper prefix of another. Prefix codes are desirable because it is easy to break a coded message into words when decoding. Figure 1, for example, shows two codes for a word set W . Code C_1 is a prefix code and C_2 is not. There is no ambiguity decoding $M = 1111100110$ according to C_1 , but decoding with C_2 produces (at least) two different source messages.

P	W	C_1	C_2
.40	not	110	11
.35	save	00	11111
.14	the	01	001
.06	trust	111	111
.05	queen	10	10

Figure 1. Two codes for the same set of source words. The first is a prefix code, the second is not.

Let us assume that the source words $W = \{w_1, \dots, w_n\}$ appear in source messages according to some fixed probability distribution $P = \{p_1, \dots, p_n\}$. For a particular code C , let $l(C, i)$ denote the length (number of bits) in $C(w_i)$. The expected word length in a random coded message is therefore $L(C) = \sum_{i=1}^n p_i \cdot l(C, i)$. Given W and P , how shall we construct an *optimal* prefix code having minimum expected word length?

Good question. Is C_1 an optimal prefix code for the probabilities given in Figure 1? Can you find a better code?

HUFFMAN CODES. In 1952 D. A. Huffman developed an elegant and efficient method for constructing optimal prefix codes given W and P . He did this by building an *encoding tree*, which is a binary tree such that every node j has an associated *cost* c_j and has either 2 or 0 children. Each source word w_i is represented by a *leaf* node i in the tree having cost assigned such that $c_i = p_i$. Left branches in encoding trees are labeled 0 and right branches are labeled 1.

Every prefix code C is represented by an encoding tree T_C . In Figure 2, for example, the encoding $C_1(\text{queen}) = 10$ is found by reading edge labels downward from the root to the leaf labeled “queen”. The prefix property is ensured because no word is an ancestor of another in the tree.

The *depth* d_i of node i is its distance from the root. The *weighted path length* of leaf node i is $a_i = p_i \cdot d_i$. The *average path length* $PL(T_C)$ of the tree is found by summing weighted path lengths over leaves and is therefore equal to the expected word length $L(C)$ of the code. In Figure 2, the “queen” node has depth 2 and weighted path length .10. The average path length for this tree is 2.46.

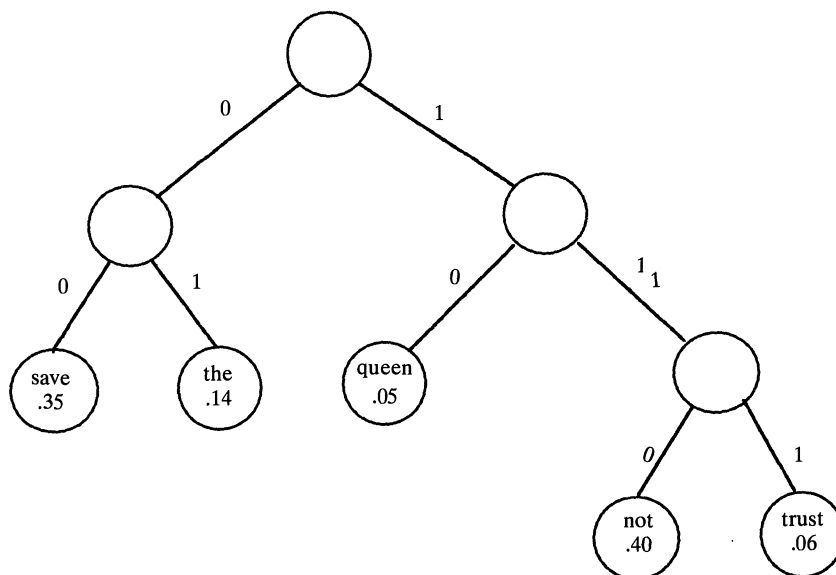


Figure 2. An encoding tree for code C_1 .

Huffman's tree construction method works as follows.

1. Begin with a list of one-node trees corresponding to words $w_1 \dots w_n$ and having costs $c_1 \dots c_n$ equal to $p_1 \dots p_n$ respectively.
2. Repeat the next two steps $n - 1$ times:
3. Find two trees Q and R in the list having smallest costs c_q and c_r at their root nodes (breaking ties arbitrarily). Remove them from the list.
4. Construct a new tree S as follows. Make a new root node s having Q as its left subtree and R as its right subtree. The cost of node s is $c_s = c_q + c_r$. Add tree S to the list.

At the end of this process the list will contain a single encoding tree, called a *Huffman tree*. We shall prove that any Huffman tree has minimal average path length.

Theorem. Let T_h be a Huffman tree constructed for a given set of words W and probabilities P . Then for any encoding tree T constructed on W and P , $PL(T_h) \leq PL(T)$.

Proof: The proof is by induction on the number of leaves in T_h . If T_h has one or two leaves the encoding tree is unique and the proof is trivial.

Suppose T_h has $n > 2$ leaves and let nodes i and j be the nodes of minimal cost that were selected in the first step of the construction. These are necessarily leaf nodes in T_h and they are necessarily siblings (having a common parent node). Construct a new tree T'_h containing $n - 1$ leaves by removing i and j : their common parent node x becomes a new leaf having cost $c_x = c_i + c_j$. Since the

only effect of this modification is to move the combined cost c_x one level closer to the root, we have $PL(T_h) = PL(T'_h) + c_x$.

Let w_x be formed by the concatenation of w_i and w_j and let $p_x = p_i + p_j$. Then T'_h is a Huffman tree constructed on the word set $W' = W - \{w_i, w_j\} + \{w_x\}$ with probabilities $P' = P - \{p_i, p_j\} + \{p_x\}$.

Now construct a tree T' from T by removing the same nodes i and j and replacing them with a new node x . Since T is an encoding tree for W these are necessarily leaf nodes, but they are not necessarily siblings in T . We have two cases:

1. If i and j are siblings then remove them and form a new leaf node from the parent x exactly as was done for T_h . The new tree T' is an encoding tree for W' and P' , and $PL(T) = PL(T') + c_x$,
2. If i and j are not siblings then adjust the tree to make them siblings: if i has greater depth than j exchange j with the sibling of i , and if j has greater depth than i exchange i with the sibling of j (if they have equal depth then it doesn't matter which gets exchanged). That is, suppose $d_i > d_j$ and that node s is the sibling of i . Detach the subtree having root s and move it (up) to j 's place in the tree, and move j (down) to s 's location in the tree. (The case $d_i < d_j$ is handled similarly.)

Moving j increases its depth by $\delta = d_i - d_j$ and increases the average path length of the tree by δc_j . But every leaf node that is moved along with s has its path length *decreased* by δ : since j was chosen to have minimal cost (except possibly for i), the net effect of the exchange operation cannot be an increase in average path length. Letting T_0 denote the new tree, we have $PL(T_0) \leq PL(T)$.

Now Case 1 holds. Remove nodes i and j and replace with x as above to form T' from T_0 . Then $PL(T') + c_x = PL(T_0) \leq PL(T)$.

By the induction hypothesis $PL(T'_h) \leq PL(T')$. Combining this with the above inequalities completes the proof. \square

OTHER CODES. One practical problem with Huffman's Code is that either the probabilities P must be estimated beforehand or the text to be encoded must be pre-scanned to determine word frequencies. This leads to inefficiencies in either the length of coded messages or in the time required to encode messages. A *dynamic* code C allows the mapping of source words to code words to change "on the fly" as the message is being encoded. Some methods (most notably Lempel-Ziv encodings) modify W dynamically as well as C . The compression factor of three mentioned earlier for Hardy's text is achieved by a dynamic method that combines several compression ideas [1].

Some codes are specialized for data other than (English) text. A digitized image of the Mona Lisa, for example, will tend to have long sequences of identical source words (which represent colors and intensities). *Run-length encoding* maps a sequence such as `yyyyyyyyybbbbbbggrrrrrrrrrr` into a sequence of pairs `9y, 6b, 2g, 10r` which may be compressed further.

For a detailed discussion of static and dynamic Huffman codes and of the Lempel-Ziv method, see Lewis and Denenberg [2]. Lelewer and Hirshberg [3] provide an extensive and detailed survey of several data compression schemes along with some experimental comparisons. Several methods for data modeling with applications to text compression are surveyed by Bell et al. [1].

REFERENCES

1. T. Bell, I. H. Witten, and J. G. Cleary, Modeling for text compression, *ACM Computing Surveys*, (21)4, December 1989, pp. 557–591.
2. H. R. Lewis and L. Denenberg, *Data Structures and Their Algorithms*, Harper-Collins 1991.
3. D. A. Lelewer and D. S. Hirschberg, Data compression, *ACM Computing Surveys*, (19)3, September 1987, pp. 261–296.

Department of Mathematics and Computer Science
P.O. Box 2239
Amherst College
Amherst, MA 01002
ccm@cs.amherst.edu

It is the consensus of opinion among college teachers of mathematics (See J. Seidlin, *Mathematics Teacher*, Dec. 1932) and science that the secondary schools produce graduates with the following general characteristics:

- (1) Worn out or weary of mathematics,
- (2) No inspiration for individual investigation,
- (3) No appreciation of accuracy,
- (4) Not able to place a decimal point in its proper place,
- (5) Direct and inverse proportions are meaningless.

—*American Mathematical Monthly*
40, (1933) p. 382

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before October 31, 1993 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10306. *Proposed by Seung-Jin Bang, Seoul, Korea.*

Find all positive integers n such that the polynomial

$$a^n(b-c) + b^n(c-a) + c^n(a-b)$$

has $a^2 + b^2 + c^2 + ab + bc + ca$ as a factor.

10307. *Proposed by John Calvin Williams, student, and I. Martin Isaacs, University of Wisconsin, Madison, WI.*

Can one construct a set \mathcal{X} of finite groups satisfying the two conditions:

- i. \mathcal{X} contains precisely one representative from each isomorphism class.
- ii. If $A \in \mathcal{X}$ is isomorphic to a subgroup of $B \in \mathcal{X}$, then A is a subgroup of B .

10308. *Proposed by Robert Connelly and John H. Hubbard, Cornell University, Ithaca, NY, and Walter Whiteley, York University, North York, Ontario, Canada.*

Suppose that $p_1, p_2, p_3, q_1, q_2, q_3$ are six points in the plane and that the distance between p_i and q_j ($i, j = 1, 2, 3$) is $i + j$. Show that the six points are collinear.

10309. Proposed by Walter Rudin, University of Wisconsin, Madison, WI.

Compute

$$\exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(A + B \cos \theta) d\theta\right)$$

when $A > B > 0$. The answer should be given as an algebraic function of A and B .

10310. Proposed by E. Rodney Canfield, University of Georgia, Athens, GA.

Fix an integer $r \geq 2$. Using Stirling's formula we may find constants c_1 and c_2 such that

$$\binom{rm}{m} \sim \frac{c_1(c_2)^m}{m^{1/2}}$$

as $m \rightarrow \infty$. Prove that the ratio $\binom{rm}{m} m^{1/2} / c_2^m$ is an increasing function of m for $m \geq 1$.

10311. Proposed by Solomon W. Golomb, University of Southern California, Los Angeles, CA.

It is well-known that if g is a primitive root modulo p , where $p > 2$ is prime, either g or $g + p$ (or both) is a primitive root modulo p^2 (indeed modulo p^k for all $k \geq 1$).

(a) Find an example of a prime $p > 2$, and a primitive root g modulo p with $1 < g < p$ such that g is *not* a primitive root modulo p^2 .

(b) Show that, among all $\phi(p-1)$ primitive roots g modulo p with $1 < g < p$, at least half of them are also primitive roots modulo p^2 .

10312. Proposed by Hongyuan Zha, IMA—University of Minnesota, Minneapolis, MN.

Let c and s be non-negative real numbers satisfying $c^2 + s^2 = 1$. Prove that, for $n > 1$,

$$s^{n-2} \sqrt{1+c}$$

is the *second* smallest singular value of the n by n upper triangular matrix

$$T_n(c) = \text{diag}(1, s, \dots, s^{n-1}) \begin{pmatrix} 1 & -c & -c & \cdots & -c \\ & 1 & -c & \cdots & -c \\ & & \ddots & \ddots & \vdots \\ & & & 1 & -c \\ & & & & 1 \end{pmatrix}.$$

10313. Proposed by O. Krafft and M. Schaefer, Rheinisch-Westfälische Technische Hochschule, Aachen, Germany.

Let $a \in [-1/5, 1)$ and let \mathcal{X}_a denote the set of random variables X satisfying $a \leq X \leq 1$. Show that

$$\max\{EX^2EX^4 - (EX^3)^2 : X \in \mathcal{X}_a\} = 2^{-6}$$

if and only if $a \in [-1/5, 1/2]$.

NOTES

Notes: (10311) The multiplicative group modulo the power of an odd prime is always cyclic, and the term *primitive root* is the traditional name in elementary number theory for a generator of this group. Fundamental properties can be found in textbooks such as I. Niven, H. S. Zuckerman and H. L. Montgomery, *An Introduction to the Theory of Numbers* (fifth edition). A consequence of (b) is that for every prime $p > 2$ there is at least one g with $1 < g < p$ which is a primitive root modulo p^k for all $k \geq 1$. **(10312)** The matrix $T_n(c)$ is a well known example in numerical linear algebra. More details can be found in G. Golub & C. Van Loan, *Matrix Computations*. It should be noted that there is no simple expression for the smallest singular value of $T_n(c)$.

SOLUTIONS

Solving the Velocity Composition Equation of Special Relativity

6659 [1991, 445]. *Proposed by Abraham Ungar, North Dakota State University, Fargo, ND.*

Let \mathbb{R}_c^3 be the subset of the Euclidean 3-space \mathbb{R}^3 given by the equation

$$\mathbb{R}_c^3 = \{\mathbf{x} \in \mathbb{R}^3: |\mathbf{x}| < c\},$$

where c is a positive constant. In the special theory of relativity c represents the speed of light, and the elements \mathbf{x} of \mathbb{R}_c^3 are *admissible velocities*. The relativistic velocity composition law is given by the equation

$$\mathbf{x} * \mathbf{y} = \frac{\mathbf{x} + \mathbf{y}}{1 + \mathbf{x} \cdot \mathbf{y}/c^2} + \frac{1}{c^2} \cdot \frac{\gamma_{\mathbf{x}}}{\gamma_{\mathbf{x}} + 1} \cdot \frac{\mathbf{x} \times (\mathbf{x} \times \mathbf{y})}{1 + \mathbf{x} \cdot \mathbf{y}/c^2}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}_c^3,$$

where $\gamma_{\mathbf{x}}$ is the *Lorentz factor*

$$\gamma_{\mathbf{x}} = \frac{1}{\sqrt{1 - \mathbf{x} \cdot \mathbf{x}/c^2}}.$$

It is known that the space \mathbb{R}_c^3 is closed under the relativistic velocity composition: if $\mathbf{x}, \mathbf{y} \in \mathbb{R}_c^3$ then $\mathbf{x} * \mathbf{y} \in \mathbb{R}_c^3$.

For given $\mathbf{a}, \mathbf{b} \in \mathbb{R}_c^3$ solve each of the two velocity composition equations

$$\mathbf{a} * \mathbf{x} = \mathbf{b} \tag{1}$$

and

$$\mathbf{x} * \mathbf{a} = \mathbf{b} \tag{2}$$

for the unknown $\mathbf{x} \in \mathbb{R}_c^3$.

Solution by Rolf Richberg, RWTH Aachen, Aachen, Germany. For $\mathbf{x} \in \mathbb{R}_c^3$, $\mathbf{x}/c \in \mathbb{R}_1^3$ and this rescaling is compatible with the definitions of $*$ in \mathbb{R}_c^3 and \mathbb{R}_1^3 , so it suffices to consider \mathbb{R}_1^3 . In this case $\gamma_{\mathbf{x}} = (1 - \mathbf{x} \cdot \mathbf{x})^{-1/2}$. Elementary vector algebra then yields that

$$\mathbf{x} * \mathbf{y} = \frac{\gamma_{\mathbf{x}}}{1 + \gamma_{\mathbf{x}}} \left(1 + \frac{1}{\gamma_{\mathbf{x}}(1 + \mathbf{x} \cdot \mathbf{y})} \right) \mathbf{x} + \frac{1}{\gamma_{\mathbf{x}}(1 + \mathbf{x} \cdot \mathbf{y})} \mathbf{y} \quad (\text{A})$$

$$\mathbf{x} \cdot (\mathbf{x} * \mathbf{y}) = 1 - \frac{1}{\gamma_{\mathbf{x}}^2(1 + \mathbf{x} \cdot \mathbf{y})} \quad (\text{B})$$

$$\gamma_{\mathbf{x} * \mathbf{y}} = \gamma_{\mathbf{x}} \gamma_{\mathbf{y}} (1 + \mathbf{x} \cdot \mathbf{y}) \quad (\text{C})$$

$$(-\mathbf{x}) * (\mathbf{x} * \mathbf{y}) = \mathbf{y} \quad (\text{D})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}_1^3$. The special case

$$\mathbf{x} * \mathbf{x} = \frac{2}{1 + \mathbf{x} \cdot \mathbf{x}} \mathbf{x}$$

of (A) is also worthy of note. It is now a simple matter to solve equation (1). From (D), we know that $\mathbf{a} * ((-\mathbf{a}) * \mathbf{b}) = \mathbf{b}$, which shows that $\mathbf{x} = (-\mathbf{a}) * \mathbf{b}$ is a solution to (1). On the other hand, (1) implies that $(-\mathbf{a}) * \mathbf{b} = (-\mathbf{a}) * (\mathbf{a} * \mathbf{x}) = \mathbf{x}$. Thus (1) has the sole solution $\mathbf{x} = (-\mathbf{a}) * \mathbf{b}$.

A slightly greater effort is required to solve equation (2). Assuming that \mathbf{x} is a solution of (2), (C) yields $\gamma_{\mathbf{x}}(1 + \mathbf{x} \cdot \mathbf{a}) = \gamma_{\mathbf{b}}/\gamma_{\mathbf{a}}$. Then (A) gives

$$\begin{aligned} \gamma_{\mathbf{b}} \mathbf{b} &= \frac{\gamma_{\mathbf{b}} \gamma_{\mathbf{x}}}{1 + \gamma_{\mathbf{x}}} \left(1 + \frac{1}{\gamma_{\mathbf{x}}(1 + \mathbf{a} \cdot \mathbf{x})} \right) \mathbf{x} + \frac{\gamma_{\mathbf{b}}}{\gamma_{\mathbf{x}}(1 + \mathbf{a} \cdot \mathbf{x})} \mathbf{a} \\ &= \frac{\gamma_{\mathbf{x}}}{1 + \gamma_{\mathbf{x}}} (\gamma_{\mathbf{b}} + \gamma_{\mathbf{a}}) \mathbf{x} + \gamma_{\mathbf{a}} \mathbf{a}. \end{aligned}$$

Now, let

$$\mathbf{v} = \frac{\gamma_{\mathbf{x}}}{1 + \gamma_{\mathbf{x}}} \mathbf{x} = \frac{\gamma_{\mathbf{b}} \mathbf{b} - \gamma_{\mathbf{a}} \mathbf{a}}{\gamma_{\mathbf{b}} + \gamma_{\mathbf{a}}}.$$

In view of $\mathbf{v} \cdot \mathbf{v} = (\gamma_{\mathbf{x}} - 1)/(\gamma_{\mathbf{x}} + 1)$, we have $\gamma_{\mathbf{x}} = (1 + \mathbf{v} \cdot \mathbf{v})/(1 - \mathbf{v} \cdot \mathbf{v})$ and

$$\mathbf{x} = \frac{2}{1 + \mathbf{v} \cdot \mathbf{v}} \mathbf{v} = \mathbf{v} * \mathbf{v}.$$

On the other hand, given $\mathbf{a}, \mathbf{b} \in \mathbb{R}_1^3$, define $\mathbf{v} = (\gamma_{\mathbf{b}} \mathbf{b} - \gamma_{\mathbf{a}} \mathbf{a})/(\gamma_{\mathbf{b}} + \gamma_{\mathbf{a}})$ and $\mathbf{x} = \mathbf{v} * \mathbf{v}$. Then, using (C), we get

$$\mathbf{v} \cdot \mathbf{v} = 1 - 2 \frac{1 + \gamma_{\mathbf{a}} * \mathbf{b}}{(\gamma_{\mathbf{b}} + \gamma_{\mathbf{a}})^2} < 1.$$

Also

$$\mathbf{x} \cdot \mathbf{x} = 1 - \left(\frac{1 - \mathbf{v} \cdot \mathbf{v}}{1 + \mathbf{v} \cdot \mathbf{v}} \right)^2 < 1$$

yields

$$\gamma_{\mathbf{x}} = \frac{1 + \mathbf{v} \cdot \mathbf{v}}{1 - \mathbf{v} \cdot \mathbf{v}} \quad \text{and} \quad \mathbf{v} = \frac{\gamma_{\mathbf{x}}}{1 + \gamma_{\mathbf{x}}} \mathbf{x}.$$

Now, by (C)

$$\begin{aligned} 2\mathbf{a} \cdot \mathbf{v} &= \frac{2}{\gamma_b + \gamma_a} \left(\gamma_b \left(\frac{\gamma_a \mathbf{a} \cdot \mathbf{b}}{\gamma_a \gamma_b} - 1 \right) - \gamma_a \left(1 - \frac{1}{\gamma_a^2} \right) \right) \\ &= \frac{2(1 + \gamma_a \mathbf{a} \cdot \mathbf{b})}{\gamma_a(\gamma_b + \gamma_a)} - 2 \\ &= \frac{\gamma_b}{\gamma_a} (1 - \mathbf{v} \cdot \mathbf{v}) - 1 - \mathbf{v} \cdot \mathbf{v}, \end{aligned}$$

and hence

$$\gamma_x(1 + \mathbf{a} \cdot \mathbf{x}) = \frac{1 + \mathbf{v} \cdot \mathbf{v}}{1 - \mathbf{v} \cdot \mathbf{v}} \left(1 + \frac{2}{1 + \mathbf{v} \cdot \mathbf{v}} \mathbf{a} \cdot \mathbf{v} \right) = \frac{\gamma_b}{\gamma_a},$$

which, with (A) yields

$$\mathbf{x} * \mathbf{a} = \left(1 + \frac{\gamma_a}{\gamma_b} \right) \mathbf{v} + \frac{\gamma_a}{\gamma_b} \mathbf{a} = \mathbf{b}.$$

This settles the case of equation (2). These formulas: $\mathbf{x} = (-\mathbf{a}) * \mathbf{b}$ in (1); and $\mathbf{x} = \mathbf{v} * \mathbf{v}$ with $\mathbf{v} = (\gamma_b \mathbf{b} - \gamma_a \mathbf{a}) / (\gamma_b + \gamma_a)$ in (2) use only expressions preserved by the mappings used to rescale c . Hence they are valid for all $c > 0$.

Editorial comment. The proposer's proof is contained in his paper, "Thomas precession and its associated grouplike structure", *Am. J. Phys.* 59 (1991), 824–834, which explores the abstract algebraic properties of addition of velocities in special relativity. In particular, weak versions of associative and commutative laws can be found which enable equations (1) and (2) to be solved by operations resembling those used in associative algebras.

Thomas N. Delmer approached the problem by analogy to the use of quaternions to study rotations in Euclidean 3-space. The matrix

$$T(\mathbf{x}) = \begin{pmatrix} \gamma_x & -\gamma_x \mathbf{x}^T / c \\ -\gamma_x \mathbf{x} / c & I + (\gamma_x - 1) \mathbf{x} \mathbf{x}^T / |\mathbf{x}|^2 \end{pmatrix}$$

describes the left action of \mathbf{x} on columns occurring as first columns of $T(\mathbf{y})$ for $\mathbf{y} \in \mathbb{R}_c^3$. The solution of equation (1) follows from the fact that $T(\mathbf{x})^{-1} = T(-\mathbf{x})$. To solve equation (2), one *linearizes* the problem by writing $T = CC^T$ where C is a matrix whose inverse is its complex conjugate \bar{C} and whose entries depend linearly on four real parameters. The equation $T\mathbf{a} = \mathbf{b}$ then takes the form $C^T \mathbf{a} = \bar{C} \mathbf{b}$, which is a system of linear equations in the parameters defining C . This use of the matrix C corresponds to the vector \mathbf{v} in the solution above.

Solved also by R. J. Chapman (U.K.), T. N. Delmer, S. Eder (student, Austria), M. Golomb, T. L. McCoy, K. McInturff, and the proposer. Two incorrect solutions were received.

An Aperiodic Sequence

E 3457 [1991, 754]. *Proposed by Herbert S. Wilf, University of Pennsylvania, Philadelphia, PA.*

Find all positive integers k such that the sequence

$$\left\{ \binom{2n}{n} \right\}_{n=0}^{\infty}$$

is periodic modulo k from some point onward.

Solution by Jerrold R. Griggs, University of South Carolina, Columbia, SC. The only such values of k are 1 and 2. The case $k = 1$ is trivial, while for $k = 2$ the familiar binomial coefficient recursion and symmetry yields

$$\binom{2n}{n} = \binom{2n-1}{n} + \binom{2n-1}{n-1} = 2\binom{2n-1}{n-1} \equiv 0 \pmod{2}$$

for all $n \geq 1$.

Next let $k = 4$. By the binomial theorem, $\binom{2n}{n} \pmod{4}$ is the coefficient of x^n in the expansion of $(1+x)^{2n}$ over $Z_4[x]$. We claim that $(1+x)^{2^m} = 1 + 2x^{2^{m-1}} + x^{2^m}$ over $Z_4[x]$ for $m \geq 1$. This is immediate for $m = 1$ and readily verified for $m > 1$ by induction on m by multiplying out $(1+x)^{2^{m+1}} = ((1+x)^{2^m})^2$. Hence $\binom{2n}{n} \equiv 2 \pmod{4}$ when $n = 2^j$ for $j \geq 0$. On the other hand, if $n \geq 3$ is not a power of 2, say $n = 2^j + r$ where $0 < r < 2^j$, we obtain over $Z_4[x]$ that

$$(1+x)^{2^n} = (1+x)^{2^{j+1}}(1+x)^{2r} = (1+2x^{2^j} + x^{2^{j+1}})(1+x)^{2r}.$$

Since $2r < n < 2^{j+1}$, the only contribution to the coefficient of x^n is $2\binom{2r}{r}$, which is divisible by 4 since $\binom{2r}{r}$ is divisible by 2. Hence $\left\{\binom{2n}{n}\right\}$ is not eventually periodic mod 4, as it contains arbitrarily long finite stretches of zeroes modulo 4.

Next suppose that k is an odd prime p . Since $p \mid \binom{p}{i}$ for all $0 < i < p$, we have $(1+x)^p = 1 + x^p$ over $Z_p[x]$. By induction, it follows for $m \geq 1$ that $(1+x)^{p^m} = 1 + x^{p^m}$. Hence for $0 \leq j < p^m$ the coefficient of x^i in $(1+x)^{p^{m+j}}$ is 1 for $i = j$ and $i = p^m$ but 0 for $j < i < p^m$. Also, the coefficient of x^{p^m} in $(1+x)^{2p^m}$ is 1 mod p . Again, the sequence $\left\{\binom{2n}{n} \pmod{p}\right\}$ contains arbitrarily long finite stretches of zeroes and cannot be eventually periodic.

Each remaining value of k is divisible by an odd prime or by 4; call this divisor d . The sequence cannot be eventually periodic mod k , else it would be eventually periodic mod d as well, which we have shown cannot happen.

Solved also by R. J. Chapman (U.K.), P. Čížek (student, France), M. Dindos (Slovakia), R. B. Eggleton (Brunei), N. J. Fine, I. Gessel, R. Holsager, I. Kastanas, K. S. Kedlaya (student), N. Komanda, O. P. Lossers (The Netherlands), D. Magagnosc, I. Nemes (Austria), A. Nijenhuis, A. Pedersen (Denmark), B. Peterson, N. G. Randolph, I. Vardi, Con Amore Problem Group (Denmark), and the proposer.

Subsets Whose Sums Are Congruent

E 3472 [1991, 956]. *Proposed by Hunter Snevily, California Institute of Technology, Pasadena, CA.*

Suppose h and k are relatively prime positive integers and $n = h + k$. Show that for each j there are $h^{-1}\binom{n-1}{k}$ k -element subsets of $\{1, 2, \dots, n-1\}$ with sum congruent to j modulo h .

Solution by Richard Holsager, American University, Washington, DC. We transform the problem slightly. For each k -element subset $A = \{a_1, \dots, a_k\}$ of $\{1, \dots, n-1\}$, labeled so that $a_1 < \dots < a_k$, define a k -element sequence $f(A) = (b_1, \dots, b_k)$ by $b_i = a_i - i$ for $1 \leq i \leq k$. Then $0 \leq b_1 \leq \dots \leq b_k \leq h-1$, and f is a bijection between the subsets and the nondecreasing k -element sequences bounded between 0 and $h-1$. Since we have reduced the sum of each set by a fixed amount $(k(k+1)/2)$, it suffices to show that the number of sequences with sum congruent to $j \pmod{h}$ is independent of j .

Consider such a sequence B . If we replace each $b_i \in B$ by $b_i + 1 \pmod h$, then we add k to the sum. If $b_k = h - 1$, then to remain in the specified set of sequences we must also replace h 's by 0's (and cyclically reorder), which does not change the sum modulo h . Since k is relatively prime to h , applying this injection h times leads us back to the original set through all the congruence classes modulo h , so each contains the same number of sequences.

Editorial comment. The proposer and the Anchorage Math Solutions Group applied a similar cyclic rotation to the original sets, viewed as subsets of an n -element set. Reiner Martin applied the properties of the q -nomial coefficient.

Solved also by G. Calinescu (Romania), R.J. Chapman (U.K.), M. Dindos (Slovakia), R. Martin (student), the Anchorage Math Solutions Group, and the proposer.

A Ratio with a Cauchy Distribution

10189 [1992, 60]. *Proposed by Ignacy I. Kotlarski, Oklahoma State University, Stillwater, OK.*

Suppose (X_1, X_2) and Y are two independent absolutely continuous random variables, where (X_1, X_2) has a distribution depending only on $X_1^2 + X_2^2$ and Y has an arbitrary distribution. Let $Z = (X_1 - X_2 Y)/(X_1 Y + X_2)$. Show that Z has a Cauchy distribution.

Solution by Kenneth Schilling, University of Michigan, Flint, MI. For $r > 0$, let $g(r)$ be the density of (X_1, X_2) at a point (x_1, x_2) with $x_1^2 + x_2^2 = r^2$. By changing to a form of polar coordinates, $X_1 = R \sin \Theta$ and $X_2 = R \cos \Theta$ with $-\pi \leq \Theta \leq \pi$, we have

$$\begin{aligned} P(\theta_1 < \Theta < \theta_2) &= \int_{\theta_1}^{\theta_2} \int_0^\infty r g(r) dr d\theta \\ &= \frac{\theta_2 - \theta_1}{2\pi}. \end{aligned}$$

Thus Θ is uniform on $(-\pi, \pi)$, so that $\tan \Theta$ is a Cauchy random variable.

Now let $\Phi = \arctan Y$ (so that Φ is a random variable on $(-\pi/2, \pi/2)$). Then $Z = \tan(\Theta + \Phi)$.

For any fixed real number ϕ , $\Theta + \phi$ is uniformly distributed modulo π . Hence, for fixed real numbers a and b ,

$$P(a < \tan(\Theta + \phi) < b) = P(a < \tan \Theta < b).$$

Since Θ and Φ are independent, we have

$$\begin{aligned} P(a < Z < b) &= P(a < \tan(\Theta + \Phi) < b) \\ &= \int_{-\pi/2}^{\pi/2} P(a < \tan(\Theta + \phi) < b) dF_\Phi(\phi) \\ &= P(a < \tan \Theta < b) \end{aligned}$$

and so Z has a Cauchy distribution.

Editorial comment. Most solvers used a similar argument, and many noted that the absolute continuity of Y is irrelevant. José Luis Palacios employed a result from B. C. Arnold and P. L. Brockett, "On distributions whose component ratios are Cauchy", *The American Statistician*, 46 (1992), 25–26, and Gérard Letac

referred to G. Letac, “Which functions preserve Cauchy laws”, *Proc. Amer. Math. Soc.*, 67 (1977), 277–286 and G. Letac, “Isotropy and sphericity: some characterizations of the normal distribution”, *Annals of Statist.*, 9 (1981), 408–417 for more general work related to this problem.

Solved also by J. A. Bucklew, D. Callan, R. J. Chapman (U.K.), S. Gleason, E. Hertz, T. Hesterberg, N. Kang (student, Korea), K. S. Kedlaya (student), G. Letac (France), A. Nijenhuis, J. L. Palacios, D. M. Rosenblum, R. Stong, Anchorage Math Solutions Group, and the proposer. Two incomplete solutions were received.

An Interval of Differences

10190 [1992, 61]. *Proposed by Peter J. Ferraro, Roselle Park, NJ.*

Suppose t is a positive integer congruent to 1 modulo 4 but not a perfect square. Put $\alpha = (1 + \sqrt{t})/2$.

(a) Prove that if n is a positive integer, then

$$1 \leq \lfloor \alpha^2 n \rfloor - \lfloor \alpha \lfloor \alpha n \rfloor \rfloor \leq \lfloor \alpha \rfloor.$$

(b) Does every integer in the interval $[1, \lfloor \alpha \rfloor]$ occur as such a difference for some positive integer n .

Solutions by John Henry Steelman, Indiana University of Pennsylvania, Indiana, PA. We prove part (a) and show that the answer to part (b) is “yes”. To get these results, let $\theta_n = \alpha n - \lfloor \alpha n \rfloor$. The one-dimensional case of Kronecker’s theorem (due to Jacobi—see J. F. Koksma, *Diophantische Approximationen*, Springer, 1936, Theorem I.5, p. 10) shows that $\{\theta_n\}$ is dense in the interval $(0, 1)$ for irrational α .

Now let t and α be as in the statement of the problem. If $t = 1 + 4r$, then a straightforward calculation yields $\alpha^2 = \alpha + r$. Thus $\alpha^2 n = \alpha n + rn$ and hence $\lfloor \alpha^2 n \rfloor = \lfloor \alpha n \rfloor + rn$. It follows that

$$(\alpha - 1)\lfloor \alpha n \rfloor = (\alpha - 1)(\alpha n - \theta_n) = rn - (\alpha - 1)\theta_n.$$

Adding $\lfloor \alpha n \rfloor$ to each side of this equation yields

$$\alpha \lfloor \alpha n \rfloor = \lfloor \alpha n \rfloor + rn - (\alpha - 1)\theta_n = \lfloor \alpha^2 n \rfloor - (\alpha - 1)\theta_n.$$

Thus we conclude that $\{\lfloor \alpha^2 n \rfloor - \alpha \lfloor \alpha n \rfloor\}$ is dense in the interval $(0, \alpha - 1)$. As $\lfloor \alpha^2 n \rfloor - \alpha \lfloor \alpha n \rfloor = \lceil \lfloor \alpha^2 n \rfloor - \alpha \lfloor \alpha n \rfloor \rceil = \lceil (\alpha - 1)\theta_n \rceil$, we see that the set of such differences consists of those integers in the interval $[1, \lfloor \alpha \rfloor]$.

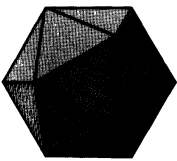
Solved also by D. Callan, R. J. Chapman (U.K.), J. Fukuta (Japan), B. Haible (Germany), R. Holzsgager, K. S. Kedlaya (student), O. P. Lossers (The Netherlands), R. Stong, B. M. M. de Weger (The Netherlands), O. Wyler, University of South Alabama Problem Group, and the proposer.

Collaborating editors: David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Jerrold R. Griggs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttmann, Frank B. Miles, Richard Pfiefer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, Edward T. H. Wang, and William E. Watkins.

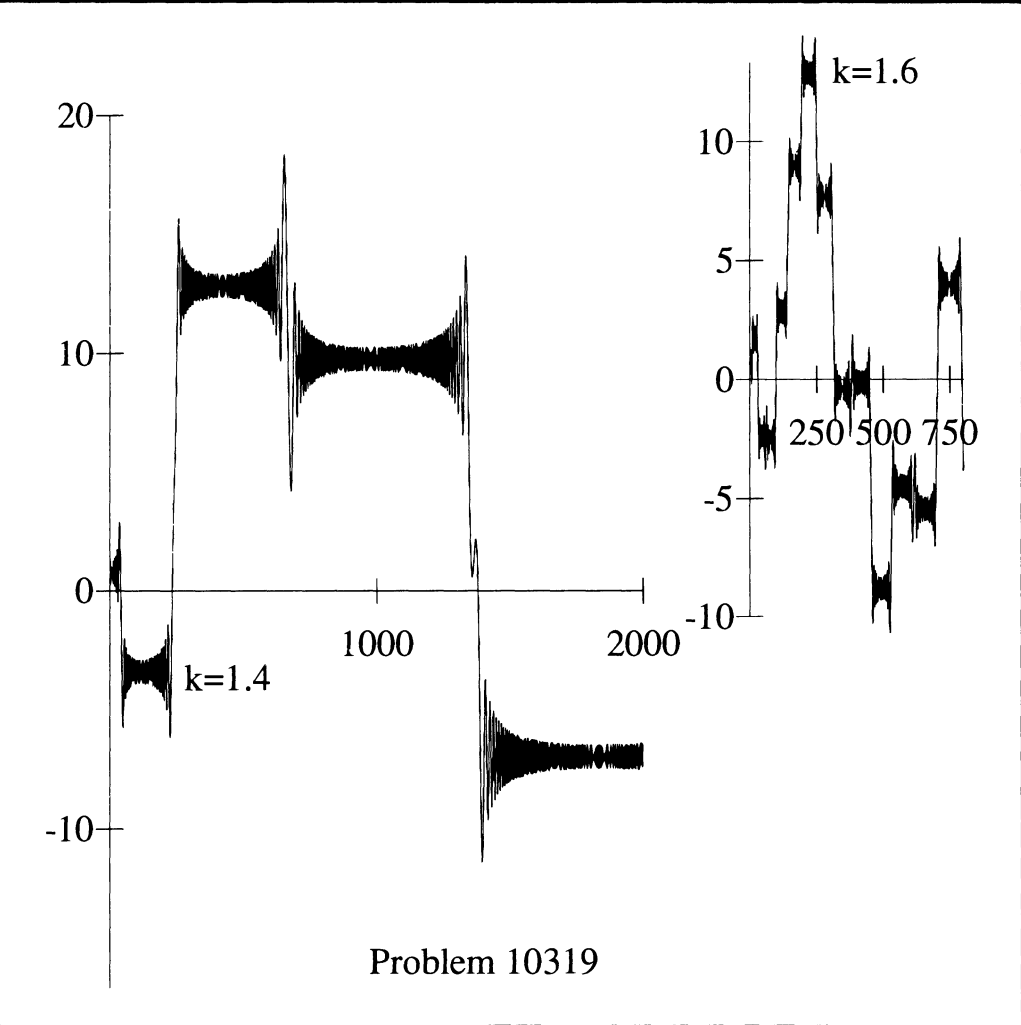
Answer to Picture Puzzle:
(p. 488)

Both Ivan Niven and Lida Barrett have been president of the MAA.

The American Mathematical Monthly



Volume 100, Number 6 / JUNE–JULY 1993



Problem 10319

NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTELEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

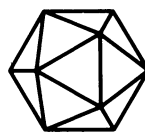
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International, Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100, Number 6 / JUNE–JULY 1993
(ISSN 0002-9890)



Contents

ARTICLES

Thomas Archer Hirst—Mathematician Xtravagent II. Student Days
in Germany / J. HELEN GARDNER and ROBIN J. WILSON 531

How to Make Wavelets / ROBERT S. STRICHARTZ 539

A Matrix Maximum / WILLIAM C. WATERHOUSE 557

Chaotic Motion of a Pendulum with Oscillatory Forcing /
S. P. HASTINGS and J. B. MCLEOD 563

An Application for the Curiosity $(\log_x N)'$ / DAVID A. WAGSTAFF,
THEODORE A. NORMAN, and DOUGLAS M. CAMPBELL 573

Vandermonde Strikes Again / MIRIAM SCHAPIRO GROSOF
and GERALDINE TAIANI 575

FEATURES

COMMENTS 530

NOTES 578

PICTURE PUZZLE 538

THE AUTHORS 583

UNSOLVED PROBLEMS

Is There a k -Anisohedral Tile for $k \geq 5$? / JOHN BERGLUND 585

PROBLEMS AND SOLUTIONS 589

REVIEWS

Mathematics in Industrial Problems. Parts 1–4.

Edited by Avner Friedman / ELLIS CUMBERBATCH 597

TELEGRAPHIC REVIEWS 600

Thomas Archer Hirst— Mathematician Xtravagant II. Student Days in Germany

J. Helen Gardner and Robin J. Wilson

Yesterday evening about 30 members of the Halifax Mechanics Institute and Mutual Improvement Society took tea together at Stott's Temperance Hotel, Broad Street, for the purpose of presenting a testimonial of respect to Mr Thomas A. Hirst, assistant to Mr Carter, land surveyor, who is about leaving the town. Mr Hirst has been an active voluntary teacher in the above society for upwards of $3\frac{1}{2}$ years, and has won the esteem and respect both of the directors and members, especially those of his own class, having taught the higher branches of mathematics with great ability.

After completing his apprenticeship on 31st August 1850, Thomas Hirst “bid adieu to surveying” forever. Remembering his earlier brief visit to Germany, and attracted by John Tyndall's enthusiasm for the University of Marburg, he resolved to study there. Tyndall was about to return to Marburg, and so Hirst went with him, arriving on 10th October after a delightful three-day journey visiting the Rhine.

Hirst quickly established for himself a daily routine which combined studying the German language with indulging his love of literature and pursuing his various scientific activities:

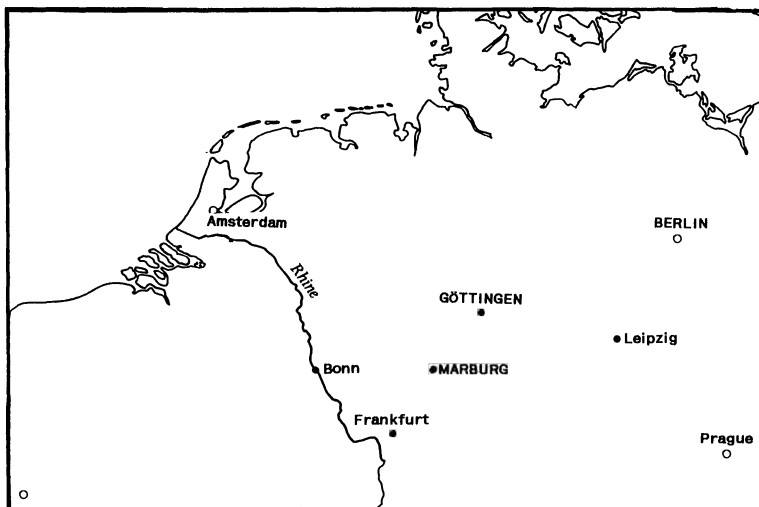
18th October 1850: . . . My time is divided thus:- I rise and get my breakfast eaten before 8 a.m. then smoke a cigar and begin the day by one of those fast earth-bound unenthusiastic essays of Montaigne. This I do medicinally to discipline myself for the practical labours of the day, then from 9 to near 1 p.m. I work in the laboratory. . . . Dinner, and afterwards German translation until dusk ($5\frac{1}{2}$), then a walk until lamp time, then German again (with an interval for tea) until $9\frac{1}{2}$; from that time to 10 my journal occupies me generally, and from 10 to 10.30 as I said Tyndall and I sweeten the day's labour with a poem.

Hirst matriculated at the University on 2nd November 1850. Being uncertain of the exact direction which his studies should take, he decided to pursue the three sciences of chemistry, physics and mathematics. His hope was that, by attending lectures in these areas, he would be able to “make choice which of the three should form the subject of my future and more particular study”.

Just as Tyndall had done previously, he attended the lectures of Robert Bunsen on chemistry, Christian Gerling on physics, and Friedrich Stegmann on mathematics. He was most enthusiastic about a laboratory session of Bunsen, but seemed rather less enthusiastic about the lectures of the other two:

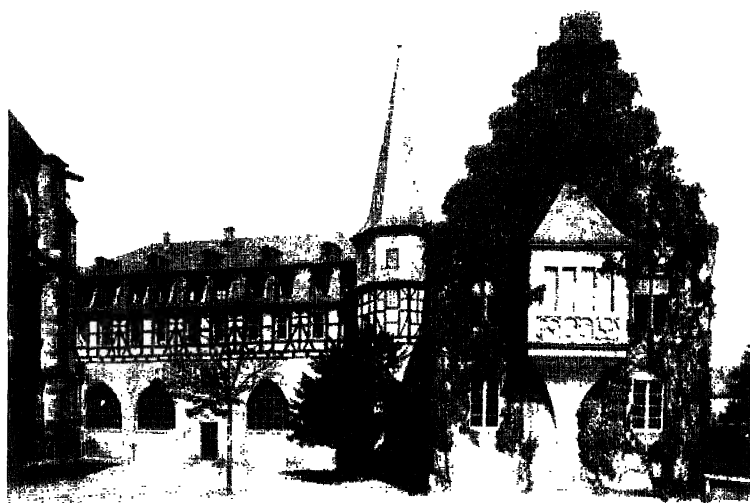
5th November 1850: . . . From 9 to 12 at laboratory. All the students began their practical course, the lectures don't commence until Thursday. Bunsen however, was present all the time and moved about from one to another, in a way that does one's heart good; able man as he is everywhere acknowledged to be, there is not the least spark of pride in him, his disposition

Map of Germany
showing Marburg, Göttingen, and Berlin



Tyndall's description of the University

Our University is not grand, it is broken into parts and presents no imposing front. Our laboratory presents rather a scoundrel-like appearance, but don't conclude hastily against it—it holds a man [Bunsen] whose superior as a chemist is not to be found within a radius of 8000 miles from the Piece Hall of Halifax. There, however, right over against me on the summit of a hill, with the sun shining upon its white walls, and its tower piercing the air, is a fine building—an astronomical observatory and physical institute, its interior furnished with costly apparatus; on the other hand I can lead you into a little room with hacked rickety benches, perhaps the whole not worth five and sixpence, where a man of genius makes his hearers forget the pooriness of his furniture, as he crushes the crust of a mathematical calculation between his fingers.



which shines through his face is a model of gentleness, geniality and integrity and humility, he is universally beloved here, and in his presence all feel at home and encouraged.

12 to 1 with Gerling. He is an old man with a good deal of the pedant about him, of weak concentrative intellect, and as is usual too much vanity. With all that however, he is what is generally called a good-hearted old boy.

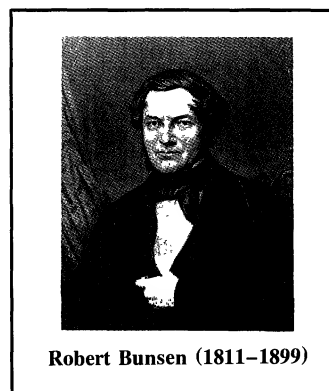
From 3 to 4 with Stegmann who differs again materially from the rest. His appearance is not prepossessing, he is an ordinary looking little man with however, a sharp nose, pale studious face, and deep sunk eyes. He bolts into the room and into his mathematics at one and the same time, wasting no time either in prelude or wordy introduction. There is a figure chalked on his black-board almost before you are aware he is present, he talks in a slow distinct voice, carves his subject deliberately piecemeal and at his exit as at his entrance you are just considering and in the middle of his last equation when you find that he has bolted, shut the door and not a vestige even of his coat flap is visible. The idea presents itself to you that if you were to follow him the moment you missed him, you would find him buried in a mathematical problem in his own room. When you come to know him, however, you find him a thorough good fellow, who always pretends less than he intends performing.

Daily his German became more fluent and his understanding more reliable. His attendance at lectures helped with this, and by mid-November Hirst began to notice the improvement himself:

19th November 1850: ... I find that with Stegmann I am learning more German than with any other. He reads mathematical operations for us to copy in writing. At first I could not copy a word, then occurred a space of time when most terrible and exasperating blanks occurred. Now only a few blanks in a page perhaps...

The turn of the year showed that Hirst had, as earlier in Halifax, quickly settled himself comfortably into a new community, and by the Spring he felt quite at home. It came as a great disappointment when Bunsen announced his impending departure from Marburg:

5th April 1851:... Bunsen called on me. He is a kind fellow indeed, during the last 2 months I have been working at the quantitative analysis of some minerals from Iceland, and he has been at great pains in explaining a theory of his as to their formation, by which theory the calculated and analysed composition shew a remarkable agreement. My analyses are a further proof of the veracity of his law, and he, thinking that some publicity would be of service and acceptable to me, proposed to me to write a small notice of my analyses and calculations for "*Liebig's Annalen*"; nay more in spite of his extreme business just now as in 2 or 3 days he leaves Marburg he has sketched out an article and to-day brought it to me. As to my share in the investigation it has been so commonplace that I should certainly refuse to publish any such article. Viewed, however, as a corroboration of *his* work, it will extend the speed of his researches and so I do it. As for the kindness to me, it was well meant, though if he knew me better he would not have offered it.



Robert Bunsen (1811–1899)

He increasingly gained satisfaction from his mathematical work, frequently to the detriment of his other subjects:

15th June 1851: ... I could do nothing well but mathematics, this week. Physics or chemistry or general literature were as arrows, that could find no entrance through my mathematical coat of mail, but glanced off merely...

Never content to relax, he worked up to sixteen hours a day, and this soon began to affect his health. Overwork and lack of physical exercise began to cause intestinal problems that were to affect him throughout his life, and he suffered an attack of dropsy. His dissatisfaction with his life style, and his frustration with his lack of progress, frequently spill over into the pages of his diaries.

27th July 1851: Many a time this week have I cursed this inward shrinking at intellectual obstacles, subjects pass by me skimmed, not penetrated into; and in spite of the day's proper number of hours having been devoted to one's task, at their close is no satisfaction. Sometimes I cry to myself, "Is it not possible to get thyself absorbed in thy work, Tom—fully?" and if not, "Is success possible?" To which I can but answer, "Thy Duty is to *do* thy work, with or without absorption therein; therefore, go about it instantly." Patience, therefore, more energetic work, and action is my need. Then from the feeling of recreation *earned*, the latter also will react on health and strength. At present my work and recreation are both accompanied with too little physical exercise. *That*, therefore, is the point to be attacked.

His life style even affected his social activities:

10th August 1851: On Tuesday evening at Museum, at a ball in the gardens. The night was chill, I dropped too suddenly from Differential Calculus into ladies' society, and could not give myself freely to the change. After an hour's unsuccessful attempt so to do, I returned, cursing the mode of life I was pursuing; next morning I had already shaken hands, however, with Diff. Calculus, and forgot the ladies...

He found relaxation in reading Tennyson and Carlyle, and translating into English the works of Goethe and Schiller:

12th October 1851: My days have been thus divided: up at 7, breakfast and Schiller until 8, then mathematics until 12.30, a walk from that time to 1, then dinner and Schiller or 'Leader' until 2.30. Once more mathematics until 5, then Physics until 7; from 7 to 8 tea and Schiller, from 8 to 11 translations of Schiller, from 11 to 12 cigar and Schiller, then to bed.

It was around this time that the direction of his future studies began to emerge more clearly. Believing that "if the heart is not in the work, there is poor chance either of success therein, or of steady perseverance", he found the idea of concentrating on mathematics increasingly compelling:

14th December 1851: ... After waverings and experiments every day brings with it the stronger conviction that my labour, in which I must find my daily discipline and duty, must be in the mathematical field. Many a time have I asked myself "what then is the absolute value of being expert at addition and subtraction? Did I come into the world to be an animated Ready-reckoner merely?" Such questions occur daily more seldom, dim visions of a higher destiny have long floated before me, as God forbid they should ever cease to do. But they *have* brought with them heretofore not merely a disturbance of the concentration necessary bravely to fulfil the day's duty, but also scattered energy to fulfil any work and even a morbid depreciation as to the value of all work itself. These dim visions of a higher density are like too full sails—dangerous, when the proportionate ballast is not there... I begin to get a gleam that there is a higher value in the multiplication table than that which teaches us that twice two make four. The Ready-reckoner even may have its transcendent side...

Life by now was incomparably better than it had been, although festivals were still

celebrated with his books, and social occasions were usually overshadowed by work:

28th December 1851: A different Christmas to any I ever spent before has again passed by. This time I had no one to share it with except Brandes *Analyt: Geom:* and Boucharlat's *Mechanics*, both which, however, if not merry, were at least interesting companions.

31st December 1851: To-night there is a ball in the Ritter. I am seated at my table at the window investigating the properties of an Ellipsoid. The music comes across the Ketzterback, mellowed by the gusts of wind—it is as if Nature had turned my room into a flute and breathed soothing harmony through it. All this serves as an accompaniment, almost unconsciously so, to my work. I have not been out for two or three days, and not the slightest idea that the New Year was on the threshold and the Old Year nearly dead: when suddenly my neighbour St Elizabeth announces the fact by tolling 12 times—simultaneously outside, where all was before in stillness, the air rings with cries of “*Prost Neu Jahr!*” ... My pen fell from my hand, and the whole past year stood before me with wondrous vividness. It has been an eventful one to me—filled with manifold new and instructive experiences. More foothold I do possess than before, so hail to thee, New Year. “Have at you,” as boxers would say.

Hirst's description of Marburg

... Marburg stands on the inner apex of an acute angle in the Lahn valley, which is a river running nearly North and South to the Rhine. Marburg stands then on the west bank, and the river flows past it with a graceful sweep into a quiescent broad hill-encircled valley to the South. It was near sunset, with a beautiful sky and a wind just strong enough to make the dying leaves sing musically and take their last and only flight high into the air before they sunk to their final rest ... Immediately before us was Augusta's Ruhe and a little farther Marburg Castle on hills of about equal height their slopes carefully terraced into rich looking gardens. To the west the sun was sinking behind the far distant purple hill between which and us was the most graceful alternation of hills and valleys with their red fallows, green, beautiful green meadows and brown woods. The spires of the church rose tapering in calm religious ascension, and the grand old castle, looked over all, with its most resigned and reverend glance. Marburg thou art indeed set in the midst of a fairy land!



By March 1852, he was completely at home in Marburg, commenting that “Germany and Germans are now to me as a native land and brothers, whereas the year

before there was ever a feeling of strangeness present". He had determined to complete his studies there during the vacation, but a letter from John Tyndall changed his plans entirely.

1st March 1852: ... An unconscious notion possessed me before that haste was needed and that it was time I left Germany. How the notion came I know not. True it is, however, when John said if he were in my place he would be in no haste the idea struck me as new, and in an hour I had made other plans, namely quite silently and unknown to any of them I will walk in and visit England, and as quickly and quietly return to Berlin or Göttingen at the beginning of next Semester. This arrangement will give me an opportunity to proceed far further with my Mathematics, and to hear some of the first German mathematicians; after which time I may sit down more confidently to a dissertation ...

However, Stegmann advised him to take his oral examination before leaving for home and then to return to complete his written dissertation. His oral examination took place on 16th March 1852, and covered physics (the motion of a pendulum, acoustics and light), crystallography and chemistry, and mathematics. It is interesting to note the areas covered by a mathematics student at that time:

...we went through part of the theory of Equations, namely the solution of general and numerical equations of higher powers—also the principal methods of elimination with two or more unknown terms. From this he turned to the theory of Curved Surfaces, principally on the Tangential Plane and Euler's Law of Curvature. He did not ask a single question in Differential or Integral Calculus, for which I was sorry. Instead of that, he asked finally a question in Descriptive Geometry, for which I was not so well prepared ... After a close examination of two hours, however, I was ordered to retire and in a few minutes was recalled, when the Decan told me they were satisfied, and that as soon as the necessary dissertation was approved I should receive my Degree ...

He could now prepare for his brief visit to England.

17th March 1852: After packing up my traps I went round to say good-bye to Professors and friends. Congratulations met me on every hand. They were mostly sincere, too, and as I had earned them I received them willingly. Shortly before dusk I took a walk towards Wertha, as I did yesterday evening before my examination. Then to prepare myself for the coming trial—now to cogitate on my past year and a half's work ... Another phase of my life is concluded, and thank God, it is an improvement on the foregoing. Here, however, it must not and shall not rest—it is but the beginning of new and better directed activities ...

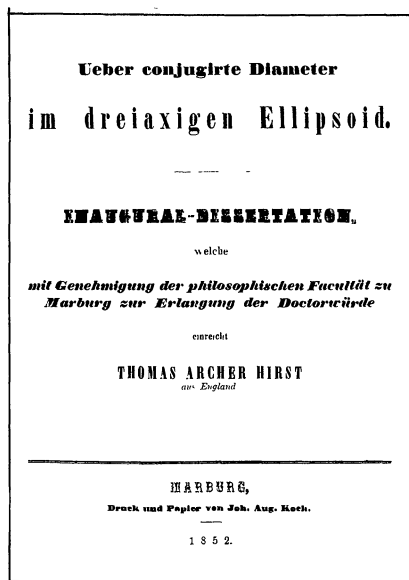
After returning to Marburg, he began work on his Ph.D. dissertation, "On conjugate diameters of the triaxial ellipsoid". By mid-June, he was able to write:

13th June 1852: ... the neck of my dissertation is already broken. Last Christmas, as I told you, I looked round the matter and spent the last part of a week thereon. Since I returned from England I have stuck pretty closely to it for ten days, and it is now done. The thing is small, it is true, but I have Stegmann and Schell's authority when I say it is a neat little investigation; both of them kindly offered to give me any assistance they could, but I did not require it ...

However, it was not all plain sailing. In particular, he had a lot of trouble trying to simplify one complicated, but important, expression. Even Wilhelm Schell, his "quick, brilliant and impulsive" supervisor, was unable to help. But a few days later, Hirst was successful:

One morning at 5 a.m. in bed a thought struck me in reference to this identical expression, to interpret whose significance had baffled me for two days. Acting on the hint, I got up, washed

myself from top to toe, and walked into it until dinner time. It was one of the luckiest hints that ever came to me—obstacle after obstacle tumbled before it, and in two days after the whole problem, much to my surprise as well as that of Stegmann and Schell, was solved. The same day Stegmann came to sit an hour with me, not having seen me for more than a week. ‘How do you get on with the Dissertation?’ he asked. ‘I think I have done it, Professor,’ I replied. He put on one of his half sarcastic, half sceptical smiles, and asked me to shew him it. I did so. He made no remark at all, until I had finished—then from his countenance one could scarcely interpret an approval—the careful dog—He then called me back very pertinently to a few important parts I had explained badly: and at length expressed his entire satisfaction saying he could suggest no improvement. I have just translated it into German roughly, and Schell is kindly correcting it for me...



Thomas Hirst's Ph.D. dissertation

“On conjugate diameters of the triaxial ellipsoid”

The dissertation was quickly approved, and Hirst was awarded his doctorate in July 1852. Like his friend Tyndall, he had completed his studies within two years, instead of the usual three.

3rd July 1852: I received orders to-day to attend upon the University Decan, Prof. Bergk, which I obeyed. It was to tell me that my dissertation had been approved of by the Philosophical Faculty, and upon delivering 120 printed copies to the University I should receive my Diploma. I took the MS. therefore, immediately to Printer Kock.

I learnt afterwards from Professor Stegmann that it first went to Prof. Gerling and his written opinion on the accompanying form was to the effect: “I find the dissertation good, and have only a few suggestions with respect to order and other trivial matters to make; I think it, however, advisable for Prof. Stegmann to certify publicly that *Mr Hirst has made it without his help*”!!! The poor old fellow, I suppose, felt slighted that after hearing his lectures on Trigonometry I declined making him my mathematical tutor. Stegmann certified accordingly that the dissertation was completed before he saw it—indeed, he might have added that he was not in Marburg when it was written.

I have received an invitation to become a member of the Mathematical Kränzchen [circle] with Professors Stegmann, Gerling, Hessel, etc. Doctors Kohlrausch, Schell, etc. to be held weekly in the open air.

11th July 1852: On Monday evening I attended the Mathematical Kränzchen in Prof. Hessel's garden. It is an interesting meeting indeed. Stegmann, with his keen intellect and quiet sarcasm, Gerling with intense vanity and essential insignificance, Hessel with his reserve and stubborn gruffness, and Schell with his unpretending, brilliant suggestions make by their contrast an interesting study...

For some time Hirst had resolved to visit Göttingen and Berlin to learn from the greatest German mathematicians of the day. In the former, he would work on magnetic experiments with Weber, and visit Gauss; in the latter, where he was to spend the winter semester, he would become a good friend of both Dirichlet and Steiner. His account of this exciting time in his life forms the topic of the next article.

ACKNOWLEDGMENTS. A typed version of the Thomas Hirst diaries is held at the Royal Institution in London, and quotations from the diaries appear here by courtesy of the Royal Institution. The diaries have been edited by W. H. Brock and R. M. MacLeod, and were published in microfiche by Mansell, London, in 1980.

*Open University
Milton Keynes MK7 6AA
England*

PICTURE PUZZLE

(from the collection of Paul Halmos)



“Is he hardier than a small forest?
(see page 596)

How To Make Wavelets

Robert S. Strichartz

§1. INTRODUCTION. The French call them *ondelettes*, these new high-tech gadgets in the arsenal of harmonic analysis. Move over, Fourier! Your series and transforms are not the only game in town. Wavelet expansions enjoy a number of good properties not available in other types of expansions. To see this in the simplest context, consider a real-valued function $f(x)$ on the interval $[0, 1]$. You can expand it in a Fourier series

$$f(x) = b_0 + \sum_1^{\infty} (b_k \cos 2\pi kx + a_k \sin 2\pi kx) \quad (1.1)$$

or you can expand it in a Haar function series

$$f(x) = c_0 + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{jk} \psi(2^j x - k) \quad (1.2)$$

where $\psi(x)$ is the function defined by

$$\psi(x) = \begin{cases} 1 & \text{if } 0 \leq x < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \leq x < 1 \\ 0 & \text{otherwise.} \end{cases} \quad (1.3)$$

(see FIGURE 1).

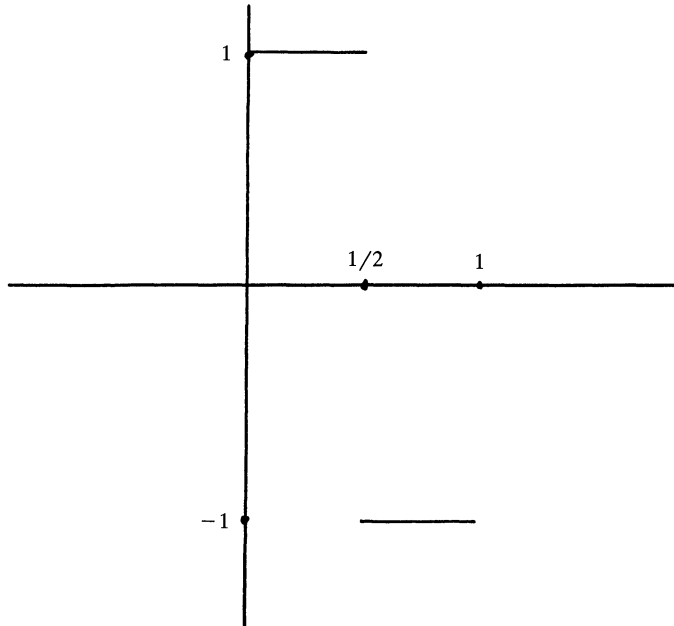


Figure 1. The graph of the generator of the Haar functions.

Both series are examples of expansions in terms of orthogonal functions in $L^2(0, 1)$. Thus there are simple formulas for the coefficients. (Exercise: Show that $\{\psi(2^j x - k)\}$ are orthogonal, but not normalized.) But the Fourier series is not well localized in space; if you are interested in the behavior of $f(x)$ on a subinterval $[a, b]$ you need to involve all the Fourier coefficients. On the other hand, the Haar series is very well localized in that to restrict attention to the subinterval $[a, b]$ you need only take the sum in (1.2) over those indices for which the interval $I_{jk} = [2^{-j}k, 2^{-j}(k+1)]$ (the support of $\psi(2^j x - k)$) intersects $[a, b]$. Furthermore, the partial sums of the Haar series (summing $0 \leq j \leq N$) clearly represents an approximation to f taking into account details on the order of magnitude 2^{-N} or greater. These two properties, *localization in space*, and *scaling*, are the hallmarks of wavelet expansions. In addition, the Haar functions are created out of a single function ψ by dyadic dilations and integer translations. Essentially the same property is shared by all the wavelet bases we will discuss, and may in fact be taken as an approximate definition of a wavelet expansion.

The wavelet expansions we are going to construct can be thought of as generalizations of the Haar series, in which the function ψ is replaced by smoother cousins. Before we can say exactly what properties we want these functions to have, and how we can go about constructing them, it is useful to backtrack and see exactly how the Haar functions arise. It will turn out to be easier if we consider the whole line as the domain of our functions.

§2. THE ROUGH-AND-READY HAAR WAVELETS. We begin with the function φ = characteristic function of the unit interval $[0, 1]$. Surely this is one of the simplest functions one can imagine, but it is chosen because it has two important properties:

(i) the translates of φ by integers, $\varphi(x - k)$, $k \in \mathbb{Z}$, form an orthonormal set of functions for $L^2(\mathbb{R})$;

(ii) φ is *self-similar*. If you cut the graph in half then each half can be expanded to recover the whole graph. This property can be expressed algebraically by the *scaling identity*

$$\varphi(x) = \varphi(2x) + \varphi(2x - 1). \quad (2.1)$$

We will call φ the *scaling* function. (In the French literature it is sometimes called “le père” and ψ is called “la mère,” but this shows a scandalous misunderstanding of human reproduction; in fact the generation of wavelets more closely resembles the reproductive life style of an amoeba.) In fact, the scaling identity essentially determines φ up to a constant multiple (exercise). The significance of the scaling identity is the following: Let V_0 denote the linear span of the functions $\varphi(x - k)$, $k \in \mathbb{Z}$ (or by abuse of notation the closure in $L^2(\mathbb{R})$ of this span, $\sum_{k=-\infty}^{\infty} a_k \varphi(x - k)$ with $\sum |a_k|^2 < \infty$). This is a natural space to consider in view of (i), since the functions $\varphi(x - k)$ form an orthonormal basis for V_0 . Of course V_0 is not all of L^2 , it is the subspace of piecewise constant functions with jump discontinuities at \mathbb{Z} . We can get a larger space by rescaling. Let $(1/2)\mathbb{Z}$ denote the lattice of half-integers $k/2$, $k \in \mathbb{Z}$, and let V_1 denote the subspace of L^2 of piecewise constant functions with jumps at $(1/2)\mathbb{Z}$. It is clear that $f(x) \in V_0$ if and only if $f(2x) \in V_1$, and the functions $2^{1/2}\varphi(2x - k)$ form an orthonormal basis for V_1 (the factor $2^{1/2}$ is thrown in to make the normalization $\|2^{1/2}\varphi(2x - k)\|_2 = 1$ hold). The scaling identity (2.1), or rather its translated version

$$\varphi(x - k) = \varphi(2x - 2k) + \varphi(2x - 2k - 1) \quad (2.1')$$

says exactly $V_0 \subseteq V_1$, since a basis for V_0 is explicitly represented as linear combinations of basis elements of V_1 . (Of course the containment $V_0 \subseteq V_1$ is clear from the description of the spaces V_0 and V_1 in terms of locations of jump discontinuities, but in the generalizations to come there will be no such simple description; however, there will be a scaling identity.)

The whole story can now be iterated, both up and down the dyadic scale. The result is an increasing sequence of subspaces V_j for $j \in \mathbb{Z}$, where V_j consists of the piecewise constant L^2 functions with jumps at $2^{-j}\mathbb{Z}$, and the functions $2^{j/2}\varphi(2^j x - k)$ for $k \in \mathbb{Z}$ form an orthonormal basis for V_j . We can pass back and forth among the space V_j by rescaling: $f(x) \in V_j$ if and only if $f(2^{k-j}x) \in V_k$, and the scaling identity (2.1), suitably rescaled, says $V_j \subseteq V_k$ if $j \leq k$. The sequence $\{V_j\}$ is an example of what is called a *multiresolution analysis*. There are two other properties of $\{V_j\}$ that are significant, namely

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad (2.2)$$

and

$$\bigcup_{j \in \mathbb{Z}} V_j \text{ is dense in } L^2 \quad (2.3)$$

(exercise).

In view of (2.3) it would seem tempting to try to combine all the orthonormal bases $\{2^{j/2}\varphi(2^j x - k)\}$ of V_j into one orthonormal basis for $L^2(\mathbb{R})$. But look, although $V_j \subseteq V_{j+1}$, the orthonormal basis $\{2^{j/2}\varphi(2^j x - k)\}$ for V_j is not contained in the orthonormal basis $\{2^{(j+1)/2}\varphi(2^{j+1}x - k)\}$ for V_{j+1} . (Indeed, there are distinct elements in the two orthonormal bases that are not orthogonal to each other.) So our first naive attempt to obtain an orthonormal basis for $L^2(\mathbb{R})$ is flawed. Can we fix it up?

Back to the drawing boards! Since $V_0 \subseteq V_1$ and we have an orthonormal basis for V_0 of the form $\{\varphi(x - k)\}$, why don't we try to complete an orthonormal basis of V_1 by adjoining functions of the form $\{\psi(x - k)\}$ for some function ψ ? This is the same thing as asking for an orthonormal basis of the desired form for the orthogonal complement of V_0 in V_1 , which we denote W_0 , so $V_1 = V_0 \oplus W_0$ (Hilbert space direct sum).

The answer is easy: we want to take ψ exactly to be the Haar function generator defined in §1. Note that ψ can be expressed in terms of φ by

$$\psi(x) = \varphi(2x) - \varphi(2x - 1) \quad (2.4)$$

which is very reminiscent of the scaling identity. Exercise: show that $\{\psi(x - k)\}$ forms an orthonormal basis for W_0 . But now we can rescale the space W_0 , so

$$V_{j+1} = V_j \oplus W_j \quad (2.5)$$

and $\{2^{j/2}\psi(2^j x - k)\}_{k \in \mathbb{Z}}$ is an orthonormal basis for W_j . If we combine conditions (2.2), (2.3) and (2.5) we obtain

$$L^2(\mathbb{R}) = \bigoplus_{j=-\infty}^{\infty} W_j \quad (2.6)$$

and since the spaces W_j are all mutually orthogonal we can now refine our naive

attempt and combine all the orthonormal bases for W_j into one grand orthonormal basis $\{2^{j/2}\psi(2^jx - k)\}_{j \in \mathbb{Z}, k \in \mathbb{Z}}$ for $L^2(\mathbb{R})$. (The only change is that we have replaced the scaling function φ by the wavelet ψ .) This gives the Haar series basis for the whole line. There is a minor variation on this theme that is perhaps more closely related to the Haar expansion on the unit interval: instead of (2.6) we can also write

$$L^2(\mathbb{R}) = V_0 \oplus \left(\bigoplus_{j=0}^{\infty} W_j \right) \quad (2.6')$$

and then combine the basis $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$ for V_0 with the bases $\{2^{j/2}\psi(2^{1/2}x - k)\}_{k \in \mathbb{Z}}$ for W_j with $j \geq 0$, to obtain an orthonormal basis for $L^2(\mathbb{R})$.

§3. MULTIREOLUTION ANALYSIS. The moral of the story so far is that we first want to build a scaling function φ and associated multiresolution analysis $\cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \cdots$ before constructing the wavelets.

Definition. A *multiresolution analysis* $\cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \cdots$ with scaling function φ is an increasing sequence of subspaces of $L^2(\mathbb{R})$ satisfying the following four conditions:

- (i) (density) $\bigcup_j V_j$ is dense in $L^2(\mathbb{R})$,
- (ii) (separation) $\bigcap_j V_j = \{0\}$,
- (iii) (scaling) $f(x) \in V_j \Leftrightarrow f(2^{-j}x) \in V_0$
- (iv) (orthonormality) $\{\varphi(x - k)\}_{k \in \mathbb{Z}}$ is an orthonormal basis for V_0 .

It follows easily from the definition that $\{2^{j/2}\varphi(2^jx - \gamma)\}_{\gamma \in \mathbb{Z}}$ forms an orthonormal basis for V_j . Since $\varphi \in V_0 \subseteq V_1$ we must have

$$\varphi(x) = \sum_{\gamma \in \mathbb{Z}} a(\gamma) \varphi(2x - \gamma) \quad (3.1)$$

for some coefficients $a(\gamma)$ satisfying

$$\sum_{\gamma \in \mathbb{Z}} |a(\gamma)|^2 = 2 \quad (3.2)$$

and in fact

$$a(\gamma) = 2 \int \varphi(x) \overline{\varphi(2x - \gamma)} dx. \quad (3.3)$$

Equation (3.1) is the analogue of (2.1), and we will refer to it as the *scaling identity*.

It follows from the definition that the scaling function determines the multiresolution analysis, but not conversely. A more difficult question is how to characterize those functions φ which are scaling functions for a multiresolution analysis. Here we expect the scaling identity to play a crucial role, but before we can say more we need to examine certain algebraic conditions on the coefficients $a(\gamma)$ that follow from the definition.

First, there is a consistency condition that arises from (iv) and (3.1). We know from (iv) that

$$\int \varphi(x - \gamma) \overline{\varphi(x)} dx = \delta(\gamma, 0) \quad (3.4)$$

(Kronecker δ). If we use (3.1) to substitute for $\varphi(x - \gamma)$ and $\overline{\varphi(x)}$ in (3.4) we

obtain

$$\begin{aligned} \sum_{\gamma' \in \mathbb{Z}} \sum_{\gamma'' \in \mathbb{Z}} a(\gamma') \overline{a(\gamma'')} \int \varphi(2x - 2\gamma - \gamma') \overline{\varphi(2x - \gamma'')} dx \\ = 2^{-1} \sum_{\gamma'' = 2\gamma + \gamma'} \sum_{\gamma' \in \mathbb{Z}} a(\gamma') \overline{a(\gamma'')} = \delta(\gamma, 0) \end{aligned}$$

after the change of variable $x \rightarrow 2^{-1}x$ and use of (3.4). We rewrite this as

$$\sum_{\gamma' \in \mathbb{Z}} a(\gamma') \overline{a(2\gamma + \gamma')} = 2\delta(\gamma, 0). \quad (3.5)$$

Note that (3.5) contains (3.2) as a special case.

Another algebraic condition arises if we assume φ is integrable and $\int \varphi(x) dx \neq 0$ (if $\int \varphi(x) dx = 0$ then the same is true for all functions in all V_j , so we would not expect to have the density condition (i)). Then we integrate (3.1) and make a change of variable to obtain

$$\begin{aligned} \int \varphi(x) dx &= \sum_{\gamma \in \mathbb{Z}} a(\gamma) \int \varphi(2x - \gamma) dx \\ &= \sum_{\gamma \in \mathbb{Z}} a(\gamma) 2^{-1} \int \varphi(x) dx \end{aligned}$$

hence

$$\sum_{\gamma \in \mathbb{Z}} a(\gamma) = 2. \quad (3.6)$$

Now we would like to reverse the procedure. *Step 1* will be to produce solutions $a(\gamma)$ to the algebraic identities (3.5) and (3.6). *Step 2* will be to define the scaling function via the scaling identity (3.1). Notice that (3.1) says that φ is a fixed point of the linear transformation

$$Sf(x) = \sum_{\gamma \in \mathbb{Z}} a(\gamma) f(2x - \gamma) \quad (3.7)$$

so it is reasonable to try to construct φ by iterating S ,

$$\varphi = \lim_{n \rightarrow \infty} S^n f \quad (3.8)$$

for some reasonable initial function f . In a later section we will discuss another method for solving (3.1). *Step 3* will be to prove that the function φ that solves (3.1) (normalized so $\|\varphi\|_2 = 1$) generates a multiresolution analysis. This is the trickiest step, because there are simple counterexamples to show that it is not always true (try $a(\gamma)$ equal to 1 for $\gamma = 0, 3$, and otherwise $a(\gamma) = 0$, and $\varphi = \chi_{[0, 3]}$, which violates (iv)). Nevertheless, many choices of $a(\gamma)$ do yield a multiresolution analysis. The difficult condition to verify is the orthonormality (iv), and we will have to postpone the discussion of when and why this holds to a later section. In Box 1 we will show how to establish the density (i) and separation (ii), given orthonormality and the additional normalization condition

$$\int \varphi(x) dx = 1. \quad (3.9)$$

Now we are ready to move on to *Step 4*, which is the construction of the wavelets themselves.

Proofs of Density and Separation

Lemma B1.1. *Let V_0 be any subspace of $L^2(\mathbb{R})$ which is contained in $L^\infty(\mathbb{R})$ and which has the property that*

$$\|f\|_\infty \leq c\|f\|_2 \quad \text{for all } f \in V_0. \quad (\text{B1.1})$$

Define V_j by the scaling condition (iii) (no assumption of the sort $V_j \subseteq V_{j+1}$ is necessary). Then (ii) holds.

Proof: The scaling condition and a simple change of variable transforms (B1.1) into

$$\|f\|_\infty \leq cm^{j/2}\|f\|_2 \quad \text{for all } f \in V_j. \quad (\text{B1.2})$$

If $f \in \cap V_j$ then (B1.2) holds for all j , and letting $j \rightarrow -\infty$ we obtain $\|f\|_\infty = 0$ hence $f = 0$. Q.E.D.

The estimate (B1.1) is easy to obtain in our case. For simplicity assume φ is bounded and has compact support, which will be the case in all our examples. Then by the orthonormality (iv) we have

$$f(x) = \sum_{\gamma \in \mathbb{Z}} \varphi(x - \gamma) \int f(y) \overline{\varphi(y - \gamma)} dy = \int K(x, y) f(y) dy$$

where $K(x, y) = \sum_{\gamma \in \mathbb{Z}} \varphi(x - \gamma) \overline{\varphi(y - \gamma)}$, so

$$|f(x)| \leq \left(\int |K(x, y)|^2 dy \right)^{1/2} \|f\|_2 = \left(\sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2 \right)^{1/2} \|f\|_2$$

and $\sum_{\gamma \in \mathbb{Z}} |\varphi(x - \gamma)|^2$ is uniformly bounded (of course much weaker conditions on φ , such as rapid decrease will also imply this).

Lemma B1.2. *Assume φ has compact support and satisfies (3.1) and (3.9), and the orthonormality condition (iv). Then the density condition (i) holds.*

Sketch of Proof: Let $P_j f(x) = 2^j \sum_{\gamma \in \mathbb{Z}} \varphi(2^j x - \gamma) \overline{f(y) \varphi(2^j y - \gamma)}$ denote the orthogonal projection onto V_j . We need to show $\lim_{j \rightarrow \infty} P_j f = f$ in L^2 for all $f \in L^2$, which is equivalent to $\lim_{j \rightarrow \infty} \|P_j f\|_2^2 = \|f\|_2^2$ by the Pythagorean theorem. It suffices to prove this for $f = \chi_A$, A any interval, by a density argument. But $\|P_j \chi_A\|_2^2 = 2^j \sum_{\gamma \in \mathbb{Z}} \int_A \varphi(2^j y - \gamma) dy^2 = 2^{-j} \sum_{\gamma \in \mathbb{Z}} \left| \int_{2^j A} \varphi(y - \gamma) dy \right|^2$. For large j , $2^j A$ will be a large interval, so essentially either $\int_{2^j A} \varphi(y - \gamma) dy = 0$ if $\gamma \notin 2^j A$ or $\int_{2^j A} \varphi(y - \gamma) dy = 1$ if $\gamma \in 2^j A$ by (3.9) (for γ in a small neighborhood of the boundary of $2^j A$ this is not quite correct, but in the limit we can ignore this detail). Thus $\|P_j \chi_A\|_2^2 \approx 2^{-j} \# \{\gamma \in 2^j A\} \approx \text{length}(A) = \|\chi_A\|_2^2$ and in the limit this becomes equality. Q.E.D.

Notice that we could essentially reverse the argument to deduce the necessity of the normalization condition (3.9).

§4. THE WAVELETS. We will consider the scaling function φ to be the first element $\varphi = \psi_0$ of a pair of functions ψ_0, ψ_1 , with ψ_1 being the wavelet generator. We would like the functions $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}, k=0,1}$ to be an orthonormal basis for V_1 . Since the functions $\{\varphi(2x - \gamma)\}_{\gamma \in \mathbb{Z}}$ already form an orthogonal basis for V_1 , the functions $\psi_0(x)$ and $\psi_1(x)$ must be linear combinations of $\varphi(2x - \gamma)$, so they must satisfy an identity

$$\psi_k(x) = \sum_{\gamma \in \mathbb{Z}} a_k(\gamma) \varphi(2x - \gamma), \quad k = 0, 1 \quad (4.1)$$

which generalizes (3.1) (of course $a_0(\gamma) = a(\gamma)$). Notice that for $k = 1$ (4.1) is an explicit formula, there is nothing to solve. But what kind of conditions should we put on the coefficients $a_k(\gamma)$? The same reasoning that led to (3.5) leads to

$$\sum_{\gamma \in \mathbb{Z}} a_j(\gamma') \overline{a_k(2\gamma + \gamma')} = 2\delta(j, k)\delta(\gamma, 0). \quad (4.2)$$

On the other hand, the condition $\int \varphi(x) dx \neq 0$ is not something we can expect to hold for ψ_1 (think of the example of Haar functions), so conditions (3.6) can only be recopied in our new notation

$$\sum_{\gamma \in \mathbb{Z}} a_0(\gamma) = 2. \quad (4.3)$$

Lemma 4.1. *If $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is an orthonormal set and if $a_j(\gamma)$ satisfy (4.2) and (4.3) then $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}, k=0,1}$ is an orthonormal set.*

Proof: It suffices to show

$$\int \psi_j(x) \overline{\psi_k(x - \gamma)} dx = \delta(j, k)\delta(\gamma, 0). \quad (4.4)$$

Now

$$\int \psi_j(x) \overline{\psi_k(x - \gamma)} dx = \sum_{\gamma' \in \mathbb{Z}} \sum_{\gamma'' \in \mathbb{Z}} a_j(\gamma') \overline{a_k(\gamma'')} \int \varphi(2x - \gamma') \overline{\varphi(2x - 2\gamma - \gamma'')} dx.$$

But the integral is $(1/2)\delta(\gamma', 2\gamma - \gamma'')$ by the orthonormality of $\varphi(x - y)$ so (4.4) reduces to (4.2). Q.E.D.

Remark. We have omitted the justification of the interchange of series and integrals, but in most of the examples we will look at the series are actually finite sums.

Thus $\{\psi_k(x - \gamma)\}_{\gamma \in \mathbb{Z}, k=0,1}$ is an orthonormal set of functions in V_1 . Is it a basis? (A kind of pseudo dimension counting argument makes this very plausible.) To show that it is a basis it suffices to represent each function $\varphi(2x - \tilde{\gamma})$ as a linear combination, and we know the coefficients will have to be

$$\begin{aligned} \int \varphi(2x - \tilde{\gamma}) \overline{\psi_k(x - \gamma)} dx &= \sum \overline{a_k(\gamma')} \int \varphi(2x - \tilde{\gamma}) \overline{\varphi(2x - 2\gamma - \gamma')} dx \\ &= \frac{1}{2} \overline{a_k(\tilde{\gamma} - 2\gamma)}. \end{aligned}$$

Thus we need to show that

$$\frac{1}{2} \sum_{k=0,1} \sum_{\gamma \in \mathbb{Z}} \overline{a_k(\tilde{\gamma} - 2\gamma)} \psi_k(x - \gamma) \quad (4.5)$$

is equal to $\varphi(2x - \tilde{\gamma})$. But if we substitute (4.1) into (4.5) we obtain

$$\sum_{\gamma \in \mathbb{Z}} \left(\frac{1}{2} \sum_{k=0,1} \sum_{\gamma' \in \mathbb{Z}} \overline{a_k(2\gamma' + \tilde{\gamma})} a_k(2\gamma' + \gamma) \right) \varphi(2x - \gamma)$$

so it suffices to show

$$\sum_{k=0,1} \sum_{\gamma' \in \mathbb{Z}} \overline{a_k(2\gamma' + \tilde{\gamma})} a_k(2\gamma' + \gamma) = 2\delta(\gamma, \tilde{\gamma}), \quad (4.6)$$

for $\tilde{\gamma} = 0$ or 1 .

Lemma 4.2. (4.6) always holds, hence $\{\psi_k(x - \gamma)\}_{\gamma \in \Gamma, k=0,1}$ is an orthonormal basis for V_1 .

Although this is a purely algebraic statement, we postpone the proof until the next section.

Theorem 4.3. Suppose φ generates a multiresolution analysis and $a_k(\gamma)$ satisfy (4.2) and (4.3) with ψ_k defined by (4.1) and $\psi_0 = \varphi$. Then the functions $\{2^{j/2}\psi_1(2^jx - \gamma)\}$ for $j \in \mathbb{Z}$, $\gamma \in \mathbb{Z}$ form an orthonormal basis of $L^2(\mathbb{R})$.

Proof: As before, let W_0 denote the orthogonal complement of V_0 in V_1 , $V_1 = V_0 \oplus W_0$. We claim $\{\psi_1(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is an orthonormal basis for W_0 . This follows because we have merely taken the basis for V_1 given by Lemma 4.2 and removed $\{\psi_0(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ which is a basis for V_0 . By scaling we obtain

$$V_{j+1} = V_j \oplus W_j$$

and

$$\{2^{j/2}\psi_1(2^jx - \gamma)\}_{\gamma \in \mathbb{Z}}$$

is an orthonormal basis for W_j . But

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$$

by the density condition.

Q.E.D.

As a simple variation on the theme, which we leave as an exercise to the reader, the set of functions $\{\varphi(x - \gamma)\}$ for $\gamma \in \mathbb{Z}$ together with $\{2^{j/2}\psi_1(2^jx - \gamma)\}$ for $j \geq 0$, $\gamma \in \mathbb{Z}$ form an orthonormal basis of $L^2(\mathbb{R})$. The advantage of this variant is that we scale only to finer and finer resolutions ($j \rightarrow +\infty$) and take care of all the coarser resolutions ($j < 0$) by the single family $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$.

In summary, we have reduced the construction of wavelets to the solution of the algebraic identities (4.2) and (4.3), modulo some technical conditions to ensure the

orthonormality condition (iv). *Step 5* will be to actually produce the solutions to (4.2) and (4.3), and *Step 6* will be to establish various properties of the wavelet functions: regularity, decay at infinity, and moment conditions.

The reason we have postponed some of the details in the construction so far is that they require a new technique. So it is now time to open the door and invite Fourier back in.

§5. THE VIEW FROM THE FOURIER TRANSFORM SIDE. Suppose we take the Fourier transform of everything in sight. Because most of our identities have a convolutional structure, we expect a simplification, with multiplicative identities arising in their place. Before doing so, let us return to the orthonormality question, because here the Fourier transform viewpoint gives us an entirely new handle on the problem. Given $\varphi \in L^2$, how can we tell from $\hat{\varphi}$ whether or not $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is orthonormal?

It will simplify matters if we adapt the convention (as in [SW]) that

$$\hat{\varphi}(x) = \int e^{2\pi i x y} \varphi(y) dy \quad (5.1)$$

so that the Fourier inversion formula is just

$$\hat{\hat{\varphi}}(x) = \varphi(-x) \quad (5.2)$$

and the Plancherel formula is

$$\|\varphi\|_2 = \|\hat{\varphi}\|_2 \quad (5.3)$$

(warning: not all the references follow this convention!).

Lemma 5.1. $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is an orthonormal set if and only if

$$\sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma)|^2 = 1 \quad \text{for all } \xi. \quad (5.4)$$

Proof: By the Plancherel formula, $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is orthonormal if and only if

$$\int e^{2\pi i \xi \gamma} |\hat{\varphi}(\xi)|^2 d\xi = \delta(\gamma, 0). \quad (5.5)$$

But the integral over \mathbb{R} can be broken up into an integral over $[0, 1]$ and a sum over \mathbb{Z} . Since $e^{2\pi i \xi \gamma}$ is periodic we obtain

$$\int_0^1 e^{2\pi i \xi \gamma} \sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma)|^2 d\xi = \delta(\gamma, 0)$$

which means that the function $\sum_{\gamma \in \mathbb{Z}} |\hat{\varphi}(\xi + \gamma)|^2$ on $[0, 1]$ has as Fourier coefficients $\delta(\gamma, 0)$, hence must be the constant function given by (5.4). Q.E.D.

Now the scaling identity (4.1) transcribes easily into the condition

$$\hat{\psi}_k(\xi) = A_k(\tfrac{1}{2}\xi) \hat{\varphi}(\tfrac{1}{2}\xi) \quad (5.6)$$

where

$$A_k(\xi) = \frac{1}{2} \sum_{\gamma \in \mathbb{Z}} a_k(\gamma) e^{2\pi i \gamma \xi} \quad (5.7)$$

(exercise, using the definition of the Fourier transform and a change of variable). Notice that $A_k(\xi)$ is smooth and periodic. Then (4.3) says

$$A_0(0) = 1 \quad (5.8)$$

and (3.9) says

$$\hat{\varphi}(0) = 1. \quad (5.9)$$

By iterating (5.6) for $k = 1$ (remember $\psi_0 = \varphi$) we obtain the infinite product representation

$$\hat{\varphi}(\xi) = \prod_{j=1}^{\infty} A_0(2^{-j}\xi) \quad (5.10)$$

(using (5.8) we can justify the local uniform convergence of the infinite product). Substituting (5.10) back into (5.6) we obtain

$$\hat{\psi}_k(\xi) = A_k\left(\frac{1}{2}\xi\right) \prod_{j=2}^{\infty} A_0(2^{-j}\xi). \quad (5.11)$$

Thus the functions A_k completely and explicitly determine the wavelets.

The most intricate part of the transcription process is the identity (4.2) that the coefficients $a_k(\gamma)$ must satisfy. What does this tell us about the functions A_k ? Rather than deal with this question directly (try it as an exercise, after the fact) we repeat the process which led to (4.2)—namely the consistency of (4.1), alias (5.6), with the orthonormality, alias (5.4). In other words, if $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is orthonormal then (5.4) must hold, and if (5.6) defines $\hat{\psi}_k$ then we want the analogue of (5.4), namely

$$\sum_{\gamma \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma) \overline{\hat{\psi}_j(\xi + \gamma)} = \delta_{jk}. \quad (5.12)$$

Now let $\eta_1 = 0$ and $\eta_2 = 1/2$. These are representations of the cosets of the subgroup \mathbb{Z} in $(1/2)\mathbb{Z}$. Then points of the lattice \mathbb{Z} can be represented uniquely as $2(\gamma + \eta_p)$ as γ varies in \mathbb{Z} and $p = 1, 2$. Then

$$\sum_{\gamma \in \mathbb{Z}} \hat{\psi}_k(\xi + \gamma) \overline{\hat{\psi}_j(\xi + \gamma)} = \sum_{p=1}^2 \sum_{\gamma \in \mathbb{Z}} \hat{\psi}_k(\xi + 2(\gamma + \eta_p)) \overline{\hat{\psi}_j(\xi + 2(\gamma + \eta_p))}$$

by the above parametrization of \mathbb{Z} , and if we substitute (5.6) and use the periodicity of A_k we obtain

$$\sum_{p=1}^2 A_k\left(\frac{1}{2}\xi + \eta_p\right) \overline{A_j\left(\frac{1}{2}\xi + \eta_p\right)} \sum_{\gamma \in \mathbb{Z}} \left| \hat{\varphi}\left(\frac{1}{2}\xi + \eta_p + \gamma\right) \right|^2.$$

The inner sum over \mathbb{Z} yields the constant 1, and so (5.12) yields the consistency condition

$$\sum_{p=1}^2 A_k(\xi + \eta_p) \overline{A_j(\xi + \eta_p)} = \delta_{jk}. \quad (5.13)$$

This is the Fourier transform equivalent of (4.2). Note that (5.13) implies

$$|A_k(\xi)| \leq 1 \quad (5.14)$$

which implies the boundedness of the Fourier transforms $\hat{\psi}_k$.

We can now easily supply the missing proof of Lemma 4.2. Notice that (5.13) says that for every ξ , the 2×2 matrix $\{A_k(\xi + \eta_p)\}$ is unitary by rows. But this is equivalent to being unitary by columns,

$$\sum_{k=0,1} A_k(\xi + \eta_p) \overline{A_k(\xi + \eta_q)} = \delta_{pq}. \quad (\text{B2.1})$$

Now substituting (5.7) into (B2.1) we obtain

$$\sum_{\gamma \in \mathbb{Z}} \left(\frac{1}{4} \sum_{k=0,1} \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma' + \gamma) \overline{a_k(\gamma')} e^{2\pi i \gamma \eta_p} e^{2\pi i \gamma'(\eta_p - \eta_q)} \right) e^{2\pi i \gamma \xi} = \delta_{pq}.$$

Regarding this as an identity between Fourier series expansions we can equate coefficients to conclude

$$\frac{1}{4} \sum_{k=0,1} \sum_{\gamma' \in \mathbb{Z}} a_k(\gamma' + \gamma) \overline{a_k(\gamma')} e^{2\pi i \gamma \eta_p} e^{2\pi i \gamma'(\eta_p - \eta_q)} = \delta_{pq} \delta(\gamma, 0).$$

Choosing $\eta_p = 0$ and summing over q we obtain (4.6) for $\tilde{\gamma} = 0$ since

$$\sum_{q=1}^2 e^{-2\pi i \gamma' \eta_q} = \begin{cases} 2 & \text{if } \gamma' \in 2\mathbb{Z} \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, choosing $\eta_p = 1/2$, multiplying by $e^{2\pi i \eta_q}$ and summing over q we obtain (4.6) for $\tilde{\gamma} = 1$.

The time has come to grasp the bull by the horns and prove the orthonormality of $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ directly. For this we will need an additional hypothesis.

Theorem 5.2. *Suppose*

$$A_0(\xi) \neq 0 \quad \text{for } |\xi| \leq \frac{1}{4}. \quad (5.15)$$

Then $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is orthonormal.

Proof: We construct a sequence of functions φ_j such that $\{\varphi_j(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is orthonormal, and such that $\varphi_j \rightarrow \varphi$ in L^2 norm as $j \rightarrow \infty$. For φ_0 we simply take $\hat{\varphi}_0(\xi) = \chi_{[-1/2, 1/2]}(\xi)$. Then $\{\varphi_0(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is orthonormal by Lemma 5.1 because (5.4) has exactly one non-zero term.

Inductively define functions φ_j by

$$\hat{\varphi}_j(\xi) = A_0\left(\frac{1}{2}\xi\right) \hat{\varphi}_{j-1}\left(\frac{1}{2}\xi\right). \quad (5.16)$$

We claim that $\{\varphi_j(x - \gamma)\}_{\gamma \in \mathbb{Z}}$ is again orthonormal. This follows immediately from (5.13) with $j = k = 0$ and Lemma 5.1. It can also be deduced from

$$\varphi_j(x) = \sum_{\gamma \in \mathbb{Z}} a_0(\gamma) \varphi_{j-1}(2x - \gamma) \quad (5.17)$$

which is the non-Fourier transform version of (5.16), and (4.2). Note that

$$\hat{\varphi}_j(\xi) = \left(\prod_{k=1}^j A_0(2^{-k}\xi) \right) \chi_{[-2^{j-1}, 2^{j-1}]}(\xi) \quad (5.18)$$

so that $\hat{\varphi}_j \rightarrow \hat{\varphi}$ pointwise, by (5.10).

We would like to show $\varphi_j \rightarrow \varphi$ in L^2 norm. This will suffice to complete the proof, because the norm limit of orthonormal sets is an orthonormal set. This is the key point of the proof, where the non-vanishing hypothesis must be used. (As an interesting exercise, see how the argument breaks down for the counterexample given in §3.)

By the Plancherel formula it suffices to show $\hat{\varphi}_j \rightarrow \hat{\varphi}$ in L^2 norm, and since we have pointwise convergence we would like to use the dominated convergence theorem. Note first that $\hat{\varphi} \in L^2$ by Fatou's theorem, since it is the pointwise limit of $\hat{\varphi}_j$ and $\|\hat{\varphi}_j\|_2 = 1$. Thus we can use a multiple of $\hat{\varphi}$ as a dominator. By comparing (5.18) and (5.10) we see

$$\hat{\varphi}_j(\xi) = \begin{cases} \frac{\hat{\varphi}(\xi)}{\hat{\varphi}(2^{-j}\xi)} & \text{if } |\xi| \leq 2^{j-1} \\ 0 & \text{otherwise.} \end{cases} \quad (5.19)$$

We claim that $\hat{\varphi}$ is bounded from below on $[-1/2, 1/2]$. The point is that $\hat{\varphi}$ is continuous, and by (5.15) $A_0(2^{-j}\xi) \neq 0$ for $|\xi| \leq 1/2$. Thus $\hat{\varphi}$ doesn't vanish on $[-1/2, 1/2]$, so $|\hat{\varphi}_j(\xi)| \leq c|\hat{\varphi}(\xi)|$ for $c = (\inf_{[-1/2, 1/2]} |\hat{\varphi}|)^{-1}$. Q.E.D.

§6. THE RECIPE. So now we have indicated all the major steps in the construction, but we have left the first to last. We need to find actual solutions to the algebraic identities (5.8), (5.13) and (5.15). There are several different approaches to this problem. We describe one that is due to Ingrid Daubechies [D1].

We look for solutions with only a finite number of $a_k(\gamma)$ different from zero, which means $A_k(\xi)$ are trigonometric polynomials. This implies that the scaling function φ and wavelet ψ_1 have compact support. This can be seen most easily from the iteration procedure (3.7) and (3.8). Say $a(\gamma) = 0$ unless $\gamma \in [0, N]$; then if f has support in $[0, N]$, so does Sf .

We concentrate first on finding the function A_0 , which must satisfy three conditions:

$$A_0(0) = 1 \quad (6.1)$$

$$|A_0(\xi)|^2 + |A_0(\xi + \frac{1}{2})|^2 = 1 \quad (6.2)$$

$$A_0(\xi) \neq 0 \quad \text{for } |\xi| \leq \frac{1}{4} \quad (6.3)$$

(here (6.1) is (5.8), (6.2) is (5.13) for $j = k = 0$ and (6.3) is (5.15)). And, of course, A_0 must be of the form

$$A_0(\xi) = \frac{1}{2} \sum_{\gamma \in \mathbb{Z}} a_0(\gamma) e^{2\pi i \gamma \xi} \quad (\text{finite sum}). \quad (6.4)$$

Note that $|A_0(\xi)|^2$ is then of the same form.

Now we already know one solution, namely

$$A_0(\xi) = \frac{1}{2}(1 + e^{2\pi i \xi}) = e^{\pi i \xi} \cos \pi \xi$$

which yields the Haar wavelets. This was deemed unsatisfactory because the wavelets are not continuous. One way to create continuity and even differentiability is to take convolution powers, or on the Fourier transform side to take ordinary powers. Thus we are tempted to try $A_0(\xi) = (e^{\pi i \xi} \cos \pi \xi)^N$ for some large N . Unfortunately (6.2) no longer holds, but we can fix this up. Note that $\cos \pi(\xi + 1/2) = -\sin \pi \xi$, so that is why $|\cos \pi \xi|^2 + |\cos \pi(\xi + 1/2)|^2 = 1$.

Now take the identity $\cos^2 \pi \xi + \sin^2 \pi \xi = 1$ and raise it to an odd power, say

$$\begin{aligned} 1 &= (\cos^2 \pi \xi + \sin^2 \pi \xi)^5 \\ &= \cos^{10} \pi \xi + 5 \cos^8 \pi \xi \sin^2 \pi \xi + 10 \cos^6 \pi \xi \sin^4 \pi \xi \\ &\quad + 10 \cos^4 \pi \xi \sin^6 \pi \xi + 5 \cos^2 \pi \xi \sin^8 \pi \xi + \sin^{10} \pi \xi. \end{aligned}$$

Take the first half of the terms for $|A_0|^2$,

$$|A_0(\xi)|^2 = \cos^{10} \pi \xi + 5 \cos^8 \pi \xi \sin^2 \pi \xi + 10 \cos^6 \pi \xi \sin^4 \pi \xi. \quad (6.5)$$

Replacing ξ by $\xi + 1/2$ turns these into the second half of the terms, so (6.2) is automatic, and (6.1) and (6.3) are easy. This gives a recipe for producing $|A_0|^2$, and it remains to take a square root of the form (6.4). We would also like to take the coefficients $a_0(\gamma)$ in (6.4) to be real, for that will yield a real-valued scaling function (and in the end real-valued wavelets as well). There is a general theorem of F. Riesz that asserts that this is possible, but in this case it is easy enough to accomplish by trial and error. Since

$$\begin{aligned} |A_0(\xi)|^2 &= \cos^6 \pi \xi (\cos^4 \pi \xi + 5 \cos^2 \pi \xi \sin^2 \pi \xi + 10 \sin^4 \pi \xi) \\ &= \cos^6 \pi \xi \left((\cos^2 \pi \xi - \sqrt{10} \sin^2 \pi \xi)^2 + (5 + 2\sqrt{10}) \cos^2 \pi \xi \sin^2 \pi \xi \right) \end{aligned}$$

we can take

$$\begin{aligned} A_0(\xi) &= (e^{\pi i \xi} \cos \pi \xi)^3 \left(\cos^2 \pi \xi - \sqrt{10} \sin^2 \pi \xi + i\sqrt{5 + 2\sqrt{10}} \cos \pi \xi \sin \pi \xi \right) \\ &= \frac{1}{8} (e^{2\pi i \xi} + 1)^3 \left(\frac{1 - \sqrt{10}}{2} + \frac{1 + \sqrt{10}}{4} (e^{2\pi i x} + e^{-2\pi i x}) \right. \\ &\quad \left. + \frac{1}{4} \sqrt{5 + 2\sqrt{10}} (e^{2\pi i x} - e^{-2\pi i x}) \right) \end{aligned} \quad (6.6)$$

which is clearly of the form (6.4) with $a_0(\gamma)$ real and $a_0(\gamma) \neq 0$ only if $-1 \leq \gamma \leq 4$.

To complete the story we need to find $A_1(\xi)$, also of the form (6.4), which satisfies

$$|A_1(\xi)|^2 + |A_1(\xi + \frac{1}{2})|^2 = 1 \quad (6.7)$$

and

$$A_0(\xi) \overline{A_1(\xi)} + A_0(\xi + \frac{1}{2}) \overline{A_1(\xi + \frac{1}{2})} = 0 \quad (6.8)$$

(these are the remaining conditions of (5.13)). Fortunately, this can be accomplished just by taking

$$A_1(\xi) = e^{2\pi i \xi} \overline{A_0(\xi + \frac{1}{2})} \quad (6.9)$$

which amounts to setting

$$a_1(\gamma) = (-1)^{\gamma+1} \overline{a_0(1-\gamma)}. \quad (6.10)$$

Then (6.7) and (6.8) follow directly from (6.2) and the periodicity of A_0 . Note also that $a_1(\gamma)$ are real valued if $a_0(\gamma)$ are.

The Fourier transform of ψ_1 is given by (5.11), which now reads

$$\hat{\psi}_1(\xi) = A_1(\frac{1}{2}\xi) \prod_{j=2}^{\infty} A_0(2^{-j}\xi) \quad (6.11)$$

with A_0 given by (6.6) and A_1 by (6.9). If we want to obtain the wavelet ψ_1 itself

rather than its Fourier transform we first find $\psi_0 = \varphi$ by iterating the mapping

$$Sf(x) = \sum_{\gamma} a_0(\gamma) f(2x - \gamma) \quad (6.12)$$

starting with any reasonable f satisfying $\int f(x) dx = 1$, and then setting

$$\psi_1(x) = \sum_{\gamma} a_1(\gamma) \varphi(2x - \gamma). \quad (6.13)$$

See FIGURES 2 and 3.

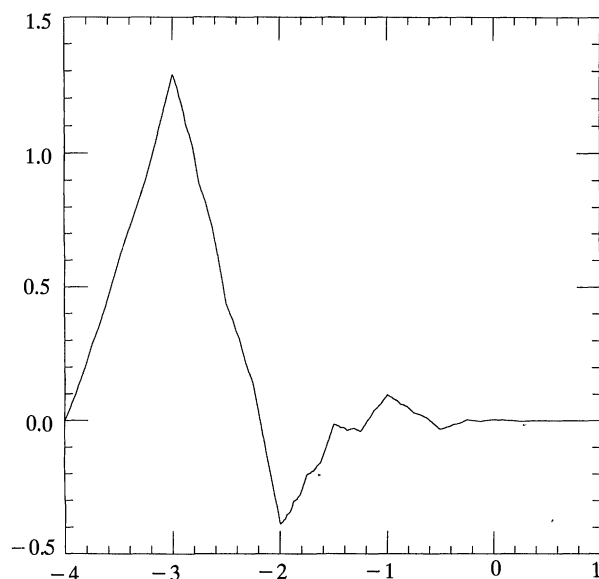


Figure 2. The graph of the scaling function φ , courtesy of David Aronstein.

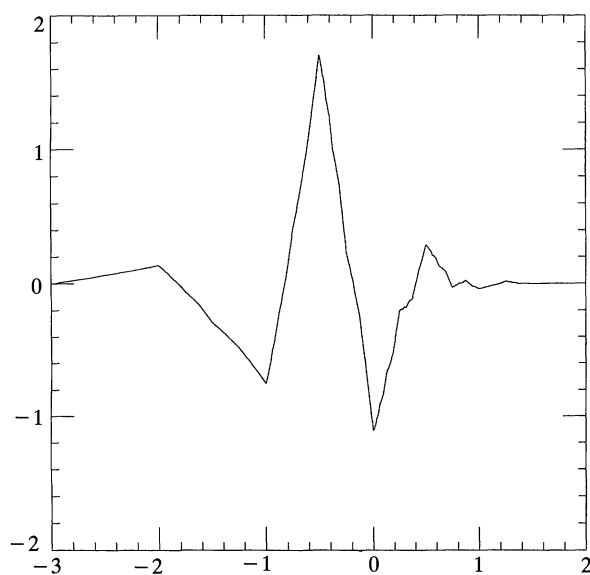


Figure 3. The graph of the wavelet generator ψ_1 , courtesy of David Aronstein.

There is an alternative approach to constructing the scaling function that yields a different wavelet basis. It has the advantage of requiring less algebra, but the disadvantage of producing wavelets that are not compactly supported. Start with the Haar basis scaling function $\chi_{[0,1]}$, whose Fourier transform is $e^{\pi i \xi} (\sin \pi \xi / \pi \xi)$, and take the N -fold convolution product

$$g = \chi_{[0,1]} * \chi_{[0,1]} * \cdots * \chi_{[0,1]} \quad (N \text{ factors})$$

so that

$$\hat{g}(\xi) = \left(e^{\pi i \xi} \frac{\sin \pi \xi}{\pi \xi} \right)^N.$$

It is easy to see that $g \in C^{N-1}$, but of course we have destroyed the orthonormality of translates by \mathbb{Z} that $\chi_{[0,1]}$ had. Too bad, but this is easily fixed. Write

$$h(\xi) = \left(\sum_{k \in \mathbb{Z}} |\hat{g}(\xi + k)|^2 \right)^{1/2}$$

and observe that h is periodic and

$$0 < c_1 \leq h(\xi) \leq c_2 < \infty.$$

Then we have only to take

$$\hat{\phi}(\xi) = \hat{g}(\xi) / h(\xi)$$

and (5.4) is automatic, so we have the orthonormality of $\{\varphi(x - \gamma)\}_{\gamma \in \mathbb{Z}}$. Notice that $\hat{g}(0) = 1$ and $\hat{g}(\gamma) = 0$ for $\gamma \neq 0$ so $\hat{\phi}(0) = 1$ as required. And it is not difficult to show that $\varphi \in C^{N-1}$.

What about the scaling identity? Well, it certainly holds for g , namely

$$\hat{g}(\xi) = B(\xi/2) \hat{g}(\xi/2)$$

where

$$B(\xi) = (e^{\pi i \xi} \cos \pi \xi)^N$$

has the required form (6.4). It then follows that

$$\hat{\phi}(\xi) = A_0(\xi/2) \hat{\phi}(\xi/2)$$

where

$$A_0(\xi) = B(\xi) h(\xi) / h(2\xi).$$

Now A_0 is periodic, so it must have the form (6.4), but the sum is no longer finite. This is where we lose the compact support of φ . On the other hand A_0 is clearly smooth, so the Fourier coefficients in (6.4) must be rapidly decreasing, which implies that φ is rapidly decreasing.

The construction of $A_1(\xi)$ and the wavelet Fourier transform $\hat{\psi}_1(\xi)$ then proceeds via (6.9) and (6.11) as before.

§7. SMOOTHNESS OF WAVELETS. How smooth are our wavelets? Since we understand them best on the Fourier transform side, we will use the principle that decay at infinity of $\hat{\varphi}$ implies smoothness of φ (we will establish smoothness of the scaling function and pass it on to the wavelets via (6.13)). For example, it is easy to show

$$|\hat{\varphi}(\xi)| \leq c(1 + |\xi|)^{-N-1-\varepsilon} \quad (7.1)$$

implies $\varphi \in C^N$. So how do we establish (7.1)?

We have the infinite product representation (5.10) which says

$$\hat{\varphi}(\xi) = \prod_{k=1}^{\infty} A_0(2^{-k}\xi) \quad (7.2)$$

and A_0 is periodic. Since each factor does not decay at infinity, why should the product? This is a mystery, which is best solved by looking at the simplest case, $A_0(\xi) = \cos \pi\xi$. Then

$$\prod_{k=1}^{\infty} \cos 2^{-k}\pi\xi = \frac{\sin \pi\xi}{\pi\xi} \quad (7.3)$$

does decay at the rate $O(|\xi|^{-1})$. (Formula (7.3) was proved by Euler, but special cases were known by Francois Viète in the late 1500's. You can prove it by considering the Fourier transform of $\chi_{[-1/2, 1/2]}$ and its scaling properties.)

Clearly, for most choices of ξ , the values of $\cos 2^{-k}\pi\xi$ will occasionally become small, and that makes the product (7.3) small. You might try to get around this by taking $\xi = 2^N$ for large N . Thus $\cos 2^{-k}\pi\xi = \pm 1$ for $k = 1, \dots, N$, so there is no decay, but then $\cos 2^{-N-1}\pi\xi = 0$ wipes you out. You can try to quantify this line of reasoning, but there is no great payoff in showing, for example, that $\sin \pi\xi/\pi\xi = O(|\xi|^{-2/3})$, so we will take (7.3) as our starting point.

The expression (6.6) for A_0 , or any of its more complicated cousins, contains $\cos \pi\xi$ as a factor, many times. Thus $\hat{\varphi}(\xi)$ contains $\sin \pi\xi/\xi$ as a factor many times, hence we expect decay. Unfortunately, the other factor grows. It is easier to work with $|A_0|^2$ given by (6.5), if we remember to take the square root at the end. We have, for the special case considered,

$$|A_0(\xi)|^2 = (\cos \pi\xi)^6 (\cos^4 \pi\xi + 5 \cos^2 \pi\xi \sin^2 \pi\xi + 10 \sin^4 \pi\xi).$$

The first factor produces decay $O(|\xi|^{-6})$. The second factor can be written $1 + 3 \sin^2 \pi\xi + 6 \sin^4 \pi\xi$ so it clearly has a maximum value 10 at $\xi = 1/2$. We can obtain a crude estimate for the growth rate produced by the second factor by the following reasoning: if $|\xi| \approx 2^N$ then there will be about N factors where $2^{-k}|\xi|$ is large, so an upper bound for the product is a constant times 10^N . But $10^N \approx |\xi|^\alpha$ for $\alpha = \log 10 / \log 2 \approx 3.32$. So the growth rate is at most $O(|\xi|^{3.32})$ so the combination gives $O(|\xi|^{-2.68})$ for $|\hat{\varphi}(\xi)|^2$ hence $O(|\xi|^{-1.34})$ for $\hat{\varphi}(\xi)$.

This is a disappointing estimate. According to (7.1) it suffices only to show that φ is continuous. It can be improved, but not by a lot. To see why, consider $\xi = 2^N/3$. Then for each of the N factors $2^{-k}\xi = 2^{N-k}/3$, $1 \leq k \leq N$, we have $1 + 3 \sin^2 2^{N-k}\pi/3 + 6 \sin^4 2^{N-k}\pi/3 = 1 + 3 \cdot (\sqrt{3}/2)^2 + 6(\sqrt{3}/2)^4 = 6.625$ so a lower bound for α is $\log 6.625 / \log 2$ which yields $O(|\xi|^{-1.636})$ as the optimal improvement.

If we consider the family of wavelets constructed as outlined in §6, we will have $|A_0(\xi)|^2$ written as the product of higher and higher powers of $\cos \pi\xi$ by more and more complicated second factors. Thus we have faster decay times faster growth in $\hat{\varphi}(\xi)$. Which wins? Well, it is a close race! It turns out that the decay wins, but the

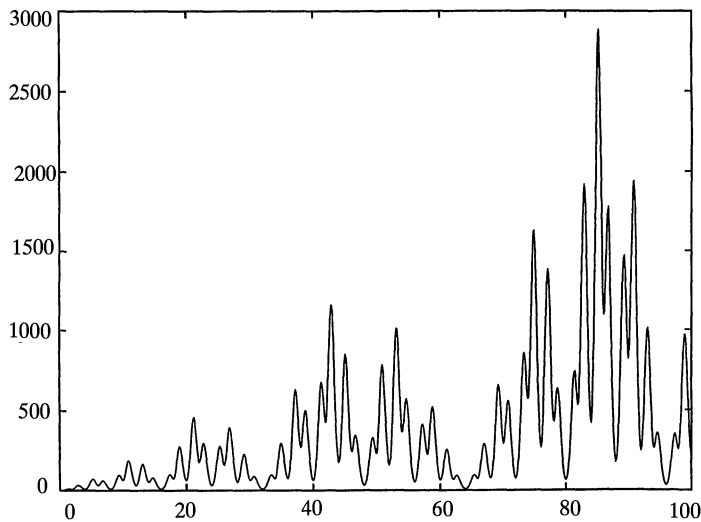


Figure 4. The graph of $\hat{\phi}$, after factoring out a power of $\sin \pi x / \pi x$, courtesy of Prem Janardhan and David Rosenblum.

crude method of estimating the growth used above is not good enough to show this. The final result ([D1], [C2]) is that to create wavelets of class C^N we need to carry out the construction starting with $(\cos^2 \pi \xi + \sin^2 \pi \xi)^M = 1$ for M on the order of $5(N + 1)$. This means that there is a rather high price to pay in terms of complexity (the algebra required to pass from $|A_0|^2$ to A_0 , for example) in order to gain a moderate amount of smoothness. (More recently, better techniques have been found to estimate the smoothness directly, without involving the Fourier transform [DL].) FIGURE 4 shows the graph of $\hat{\phi}(\xi)$. See [JRS] for a discussion of the surprising self-similarity properties of this function.

In addition to smoothness, another important property of wavelets is the vanishing moment conditions

$$\int_{-\infty}^{\infty} x^k \psi_1(x) dx = 0, \quad k = 0, 1, \dots, N \quad (7.4)$$

which are equivalent to the vanishing of the Fourier transform to high order at the origin,

$$\left(\frac{d}{d\xi}\right)^k \hat{\psi}_1(0) = 0, \quad k = 0, 1, \dots, N. \quad (7.5)$$

In contrast to smoothness, however, it is only the wavelet, not the scaling function, which enjoys this property. The significance of this condition is that it implies a weak form of localization in the frequency (Fourier transform) variable, since the Fourier transform of $\psi_1(2^j x - k)$ is mainly concentrated around values of $|\xi|$ on the order of 2^j . (There is yet another family of wavelets in which the Fourier transform is actually supported in an annular region $c_1 2^j \leq |\xi| \leq c_2 2^j$. See [M] for a description of these “Littlewood-Paley” type wavelets.) For our wavelets the verification of (7.5) is easy. From (6.11) we see that $\hat{\psi}_1$ has a factor $A_1((1/2)\xi)$, and from (6.9) we see that A_1 at $\xi = 0$ has the same order zero as A_0 at $\xi = 1/2$. But A_0 has a factor of $\cos \pi \xi$ to a power, hence vanishes at $\xi = 1/2$ to order 3 in our particular example, and to order M if we start with $(\cos^2 \pi x + \sin^2 \pi x)^M = 1$ in our construction. Note that in general conditions (6.1) and (6.2) imply that

$A_0(1/2) = 0$, and the flatter we make A_0 near $\xi = 0$, the more it vanishes near $\xi = 1/2$.

§8. CONCLUDING REMARKS. Why not try to create your own designer wavelets by programming the recipe given in §6, and taking the square root of $|A_0(\xi)|^2$ in a different way? For a more detailed discussion of the Riesz Lemma for doing this see [D1].

For further information about wavelets, including historic accounts and attribution of results, see the books [M], [BF], [BC] or the expository lectures [D2] and [FJW]. The term “wavelet” is also used to describe expansions in terms of functions which are not orthogonal. These wavelets have a simpler algebraic description, which is useful for some applications. An expanded version of this article, including a discussion of wavelet bases in several variables, will appear in [BF]. None of the theorems or proofs presented here are original; I have only tried to organize the material in a way that is easy to digest.

REFERENCES

-
- [BF] J. Benedetto and M. Frazier (editors), “Wavelets: Mathematics and Applications,” CRC Press, to appear.
 - [BC] G. Beylkin, R. Coifman, I. Daubechies, S. Mallet, Y. Mayer, M. B. Ruskai and L. Raphael (editors), *Wavelets and Their Applications*, Jones and Bartlett, 1992.
 - [C1] A. Cohen, *Ondelettes, analyses multi-résolutions et filtres miroir en quadrature*, Annales de l’Institut Henri Poincaré, Analyse non linéaires 7 (1990), 439–459.
 - [C2] A. Cohen, *Construction de bases d’ondelettes α -Hölderiennes*, Rev. Mat. Iberoamericana 6 (1990), 91–108.
 - [D1] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math. 41 (1988), 909–996.
 - [D2] I. Daubechies, *Ten Lectures on Wavelets*, Notes for the CBMS Conference (Lowell), SIAM 1992.
 - [DL] I. Daubechies and J. Lagarias, *Two-scale difference equations, I and II*, SIAM J. Math. Anal. 22 (1991), 1388–1410, 23 (1992), 1031–1079.
 - [FJW] M. Frazier, B. Jawerth and G. Weiss, *Littlewood-Paley theory and the study of functions spaces*, CBMS, Regional Conference Series in Mathematics, no. 79, American Mathematical Society, Providence, RI, 1991.
 - [JRS] P. Janardhan, D. Rosenblum and R. S. Strichartz, *Numerical experiments in Fourier asymptotics of Cantor measures and wavelets*, to appear, *Exper. Math.*.
 - [M] Y. Meyer, *Ondelettes et Opérateurs*, 2 volumes, Hermann, Paris, 1990.
 - [SW] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

*Mathematics Department
Cornell University
Ithaca, NY, 14853*

A Matrix Maximum

William C. Waterhouse

1. INTRODUCTION. In a recent paper [KZ], Kwong and Zettl described the solution to a maximization problem involving a 2 by 2 matrix A with real entries. They used a special way of normalizing the matrix, and it led them into computations that could only be done as symbolic manipulations on a computer. I want to show how a more geometric normalization will uncover a latent symmetry in the problem and thereby reduce the computation to comprehensible steps.

To understand the normalization, we should begin with a simpler, more familiar question: what is the maximum of $\|Ax\|/\|x\|$, where $\|x\|$ denotes length in the plane? The first thing to observe is that scaling x does not change the ratio, and thus we only have to consider its values for x on the unit circle. Now A clearly must map the unit circle $x_1^2 + x_2^2 = 1$ to an ellipse (or straight line segment, if A is singular). Thus if the semi-axes of the ellipse are (say) k and m , then the larger of those two is the maximum of $\|Ax\|/\|x\|$.

We could turn this geometric analysis into a computation of the maximum, but it is more important to see how it leads to an expression for the structure of A (see Figure 1). Let $R(\phi)$ be the matrix of rotation by angle ϕ , so

$$R(\phi) = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}.$$

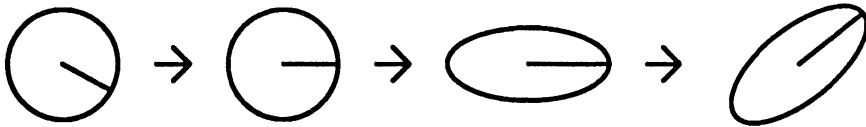


Figure 1. Structure of a Linear Mapping

The inverse of $R(\phi)$ is of course $R(-\phi)$. Take one of the half-axes of the ellipse, say of length k , and suppose it makes an angle α with the positive x -axis. Then $R(-\alpha)A$ maps the unit circle to an ellipse with axes along the coordinate axes. The half-axis lengths are still k and m , and so we can multiply by a diagonal matrix to get $\text{diag}(k, m)^{-1}R(-\alpha)A$ mapping the unit circle to itself. Hence it is an orthogonal mapping, either some rotation $R(\beta)$ or $\text{diag}(1, -1)$ times $R(\beta)$. Multiplying through and absorbing any negative sign into m , we obtain the following result, a variant of the “singular value decomposition.”

Theorem 0. Every 2 by 2 matrix A can be written in the form

$$A = R(\alpha) \begin{pmatrix} k & 0 \\ 0 & m \end{pmatrix} R(\beta)$$

for some angles α, β and some constants k, m .

You can verify that this geometric proof [OG, p. 343–6] does indeed have a modification that works when A is singular. The same idea can be stated purely in terms of linear algebra: the first step is to observe that AA^T is symmetric with positive eigenvalues, and hence it equals $R(\alpha)\text{diag}(k^2, m^2)R(-\alpha)$ for suitable k, m, α . Then if you take $B = R(\alpha)\text{diag}(k, m)R(-\alpha)$, you can easily verify that $B^{-1}A$ is orthogonal. (The theorem is actually valid in any number of variables, with special orthogonal matrices in place of rotations. See [H, p. 169] or [G, p. 286].) For our purposes, the advantage of this decomposition of A is that it incorporates information about the relation between $\|Ax\|$ and $\|x\|$. Observe that we have our choice of the order in which the diagonal entries occur, and so for nonzero A we can always suppose that k is nonzero.

2. NORMALIZING THE PROBLEM. Fix now an invertible A . The problem solved by Kwong and Zettl is to find the maximum of

$$\frac{\|Ax\|^2}{\|x\| \cdot \|A^2x\|}$$

for nonzero x . As A is invertible, we can make a change of variable to replace x by Ax ; thus it is equivalent to say that we want the maximum of

$$\frac{\|x\|^2}{\|A^{-1}x\| \cdot \|Ax\|}.$$

Clearly the ratio is again homogeneous in x , so the maximum can be found on the unit circle. As Kwong and Zettl observed, we have

$$\|Ax\| = \|R(\phi)Ax\| = \|R(\phi)AR(-\phi)[R(\phi)x]\|,$$

and similarly for A^2 , so the maximum for A is the same as for $R(\phi)AR(-\phi)$. Furthermore, in this problem there is a homogeneity in A , so the maximum does not change if we multiply A by a nonzero scalar. Now the decomposition of the previous section shows us a nice way to simplify A using these operations: we can first conjugate by a rotation to cancel the factor $R(\alpha)$, and then we can multiply by a scalar to make the first entry in the diagonal factor equal to 1. Thus we have a promising normalization:

Theorem 1. Let A be any 2 by 2 matrix not identically zero. Multiplying by a scalar and conjugating by a rotation, we can reduce A to the form $MR(\theta)$, where θ is some angle and $M = \text{diag}(1, m)$ for some constant m . \square

We now take A to be of the form $MR(\theta)$, and we let $K(m, \theta)$ be the maximum as x varies over the unit circle. To eliminate square roots, we can work with the square of the ratio and compute $K^2(m, \theta)$. Let t parametrize the unit circle by angle, so $x(t)$ will have entries $\cos(t)$ and $\sin(t)$. For brevity, set

$$Q(t) = \|Mx(t)\|^2 = \cos^2(t) + m^2 \sin^2(t) = m^2 + (1 - m^2)\cos^2(t). \quad (1)$$

Obviously $R(\theta)x(t) = x(t + \theta)$, so

$$\|Ax(t)\|^2 = \|MR(\theta)x(t)\|^2 = Q(t + \theta).$$

Similarly, we have

$$\|A^{-1}x(t)\|^2 = \|R(-\theta)M^{-1}x(t)\|^2 = \|M^{-1}x(t)\|^2 = Q(\pi/2 - t)/m^2.$$

Thus we have:

Lemma 2. Let $f(t) = m^2/Q(t + \theta)Q(\pi/2 - t)$. Then $K^2(m, \theta) = \max_t f(t)$. \square

Before going on, it might be good to look at the graph of $f(t)$ in a few examples. The two basic types are illustrated in Figures 2 and 3. All use the same angle $\theta = \pi/4$, so they also illustrate the change of behavior with m . Observe that there are certain values of t depending only on θ (in this case, $t = \pi/8$ and $t = \pi/2 +$

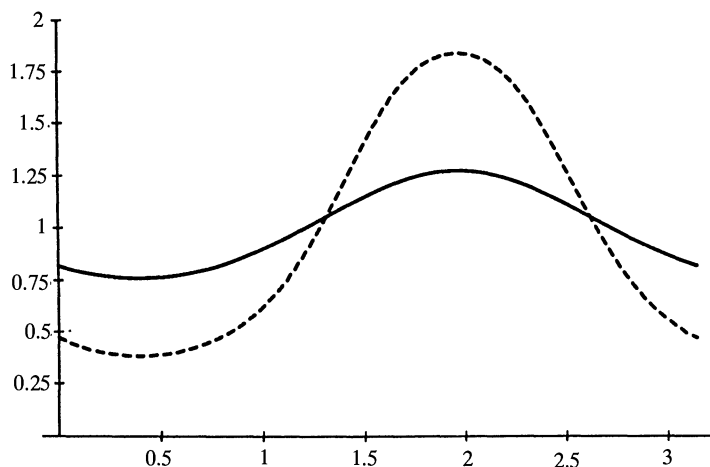


Figure 2. $f(t)$ with $\theta = \pi/4$ and $m = 1.2$ (solid), $m = 1.8$ (dashed)

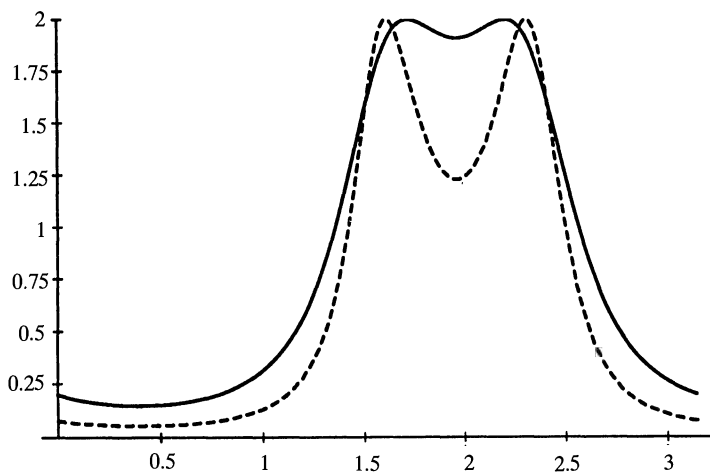


Figure 3. $f(t)$ with $\theta = \pi/4$ and $m = 3$ (solid), $m = 5$ (dashed)

$\pi/8$) that always give local extrema. For large m , however, the absolute maximum occurs elsewhere and has a value independent of m . We shall show that these properties are true in general.

3. THE LATENT SYMMETRY. It is clear from the original homogeneity that $f(t + \pi) = f(t)$. But the expression for f in terms of Q shows that there is also a latent symmetry. To bring it out, we set $p = (\pi/2 + \theta)/2$ and $u = t - (\pi/2 - \theta)/2$; then we get $f(t) = m^2/g(u)$ with

$$g(u) = Q(p + u)Q(p - u). \quad (2)$$

When we expand, we get

$$\begin{aligned} Q(p + u) &= m^2 + (1 - m^2)[\cos(p)\cos(u) - \sin(p)\sin(u)]^2 \\ &= m^2 + (1 - m^2)[\cos^2(p)\cos^2(u) + (1 - \cos^2(p))(1 - \cos^2(u))] \\ &\quad - 2(1 - m^2)\cos(p)\cos(u)\sin(p)\sin(u). \end{aligned} \quad (3)$$

The symmetry now lets us observe that $Q(p - u)$ is the same as $Q(p + u)$ except for the sign of the last term, which will be reversed. Thus the product $g(u)$ will be the difference of the squares. We can see at once that sines and cosines will occur in g only as squares, and thus we are going to be able to use cosines alone; furthermore, we see that the result will be quadratic in the variable $W = \cos^2(u)$. To find it explicitly, we have to do some straightforward computation. Of course you can save time and avoid mistakes by doing it on a computer, but it can certainly be handled by hand. Here is the result.

Lemma 3. Set $W = \cos^2(u)$. Then $g(u) = G(W)$ where

$$\begin{aligned} G(W) &= (1 - m^2)^2 W^2 + 2(1 - m^2)[m^2 \cos^2(p) + \cos^2(p) - 1]W \\ &\quad + (1 - m^2)^2 \cos^4(p) - 2(1 - m^2)\cos^2(p) + 1. \quad \square \end{aligned}$$

4. COMPUTATION OF THE CRITICAL VALUES. Our $G(W)$ is identically 1 if $m^2 = 1$; otherwise it is quadratic. The square term is positive, and hence the unique extreme value of G will be its absolute minimum. There is basically no difficulty in finding its extremum; again the computation takes a little work, but it is not hard. The appearance of the factor $(1 - m^2)$ to appropriate powers in the coefficients helps make the result particularly nice:

Lemma 4. For $m^2 \neq 1$, the function $G(W)$ has a unique extremum (a minimum) at the point

$$W = \frac{1 - \cos^2(p) - m^2 \cos^2(p)}{1 - m^2}. \quad (4)$$

The value at that point is $4m^2 \cos^2(p)(1 - \cos^2(p))$. \square

Now we can begin to translate this result back into our original variables. We had $g(u) = G(\cos^2(u))$, and hence we have

$$g'(u) = -2\cos(u)\sin(u)G'(\cos^2(u)).$$

Thus the critical points of $g(u)$ occur at points where $\cos^2(u)$ is either 0 or 1 or the W in (4) where G has its minimum. When $\cos^2(u)$ is either 0 or 1, we can see that the last term in (3) is zero, and so we can evaluate $g(u) = Q(p + u)Q(p - u)$

directly; we get the values

$$\left[m^2 + (1 - m^2)\cos^2(p) \right]^2 \quad \text{and} \quad \left[m^2 + (1 - m^2)\sin^2(p) \right]^2.$$

We have $p = (\pi/2 + \theta)/2$, and hence

$$2\cos^2(p) - 1 = \cos(2p) = \cos(\pi/2 + \theta) = -\sin(\theta).$$

Thus $\cos^2(p) = (1 - \sin(\theta))/2$, and similarly $\sin^2(p) = (1 + \sin(\theta))/2$. The values at the first two types of critical points then come out to be

$$\left(\frac{m^2 + 1 \pm (1 - m^2)\sin(\theta)}{2} \right)^2. \quad (5)$$

The value at the minimum of G is even simpler; it comes out to be just $m^2 \cos^2(\theta)$.

Of course it is possible that the G -minimum occurs at a point that cannot be a value of $\cos^2(u)$. The reduction from p to θ shows that this minimum occurs when

$$W = \frac{1 - m^2 + (1 + m^2)\sin(\theta)}{2(1 - m^2)}.$$

Hence we need to determine when this value is between 0 and 1. That is just a two-line computation, yielding the condition

$$\left| \left(\frac{1 + m^2}{1 - m^2} \right) \sin(\theta) \right| \leq 1. \quad (6)$$

Thus we have finished our computations, which we can summarize quite briefly:

Theorem 5. *The function $g(u)$ has the values (5) at the critical points where u is a multiple of $\pi/2$. If (6) is false, these are the only critical points; but if (6) is true, g also has an absolute minimum with value $m^2 \cos^2(\theta)$. \square*

5. THE MAIN THEOREM. If we multiply by the denominator, we can state the basic inequality in a way that makes sense for singular matrices, and (like [KZ]) we include that in our final version of the result.

Theorem 6. *Let A be a nonzero 2 by 2 matrix, and let scaling and conjugation by rotations reduce A to the form $MR(\theta)$, where $M = \text{diag}(1, m)$. Let K denote the smallest constant (if any) that makes $\|Ax\|^2 \leq K \cdot \|x\| \cdot \|A^2x\|$ true for every vector x . If*

$$\left| \left(\frac{1 + m^2}{1 - m^2} \right) \sin(\theta) \right| \leq 1,$$

then $K = 1/|\cos(\theta)|$. When $m = 0$ and $\cos(\theta) = 0$, the inequality does not hold with any constant. In all other cases, K is the larger of the two values

$$\frac{2|m|}{m^2 + 1 \pm (1 - m^2)\sin(\theta)}.$$

Proof: When $m \neq 0$, this theorem follows at once from (2), Lemma 2, and Theorem 5. (The case $m^2 = 1$ was excluded in Section 4, but the theorem gives the correct answer in that case, as (6) is then not true.) For $m = 0$, we can hold θ fixed and let m approach 0; as we are working with the compact set of vectors of norm 1, the constant K in the limiting case will be the limit of the approximating ones. If

$\sin(\theta) \neq \pm 1$, then the expression in (6) is less than 1 for all m close to 0; the extreme is thus equal to $1/|\cos(\theta)|$, as the theorem says. If $\sin(\theta) = \pm 1$, then we are in the last case for all small m ; the maximum there is $1/|m|$, which goes to infinity as m goes to 0. \square

The case where no bound exists corresponds to the nilpotent normalized matrix

$$\begin{pmatrix} 0 & \pm 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{pmatrix}.$$

In the other cases, we know the u giving the maximum, and so we could trace back the normalizations to compute the angle of maximum for the original A .

6. RELATION TO THE EARLIER TREATMENT. In [KZ], the matrices were normalized to have their diagonal entries both equal to 1 or both equal to 0. Consequently, the results there look rather different, and I want to conclude by displaying the connection with the formulas derived here. First we rotate: if we set $\psi = (\pi/2 - \theta)/2$, then it is easy to check that

$$R(-\psi)MR(\psi) = \frac{m+1}{2} \begin{pmatrix} \cos(\theta) & \frac{m-1}{m+1} - \sin(\theta) \\ \frac{m-1}{m+1} + \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

To avoid special cases and sign distinctions, let us suppose that $m > 1$ and both $\sin(\theta)$ and $\cos(\theta)$ are positive. We can then scale to get

$$\begin{pmatrix} 1 & b \\ c & 1 \end{pmatrix}$$

with

$$b = \frac{1}{\cos(\theta)} \left(\frac{m-1}{m+1} - \sin(\theta) \right) \quad \text{and} \quad c = \frac{1}{\cos(\theta)} \left(\frac{m-1}{m+1} + \sin(\theta) \right).$$

Using the notation of [KZ], we introduce $h = c - b$ and $r = (1 + h^2/4)^{1/2}$. Their assertion is that K for this matrix is $|1 - bc|/(1 + b^2)$ except when b is between $2(1 - r)/h$ and $2(1 + r)/h$, in which case $K = r$. It is easy to check that (in our notation) $h = 2 \tan(\theta)$ and $r = 1/\cos(\theta)$, while $|1 - bc|/(1 + b^2)$ comes out to be

$$\frac{2m}{m^2 + 1 + (1 - m^2)\sin(\theta)}.$$

Readers might find it a pleasant exercise to check that the betweenness condition on b is equivalent to our condition (6).

REFERENCES

-
- [G] F. R. Gantmacher, *The Theory of Matrices*, trans. K. A. Hirsch, Vol. I, Chelsea, 1959.
[H] P. R. Halmos, *Finite-Dimensional Vector Spaces*, Van Nostrand, 1958.
[KZ] M. K. Kwong and A. Zettl, Norm inequalities for the powers of a matrix, *Amer. Math. Monthly*, 98 (1991), 533–538.
[OG] W. F. Osgood and W. C. Graustein, *Plane and Solid Analytic Geometry*, Macmillan, 1920.

Department of Mathematics
The Pennsylvania State University
University Park, PA 16802

Chaotic Motion of a Pendulum with Oscillatory Forcing

S. P. Hastings and J. B. McLeod

I. INTRODUCTION. The mathematical theory of “chaos” has grown rapidly in the last twenty years, with one landmark being the 1975 paper [9] of Y. Li and J. Yorke which appeared in this journal. Indeed, we understand that this paper included the first use of the word chaos in the context of dynamical systems. The subject dominates dynamical systems theory today, in the literature of both mathematics and physics. This is despite a lack of unanimity on what sorts of behavior should be called chaotic, or any firm definition of associated concepts, such as “strange attractor,” or “sensitivity to initial conditions.”

The Li-Yorke paper, which turned out to be a rediscovery of some of the results of the Soviet mathematician A. N. Sharkovsky eleven years earlier [3, 13], dealt with iterations of maps of an interval into itself. Even today, chaos theory is far more developed in the case of maps, in one or two dimensions particularly, than it is for smooth dynamical systems, such as differential equations. This is largely because differential equations are so much harder. For example, the famous set of differential equations found by E. N. Lorenz [10] in the context of meteorological investigations is still not understood very well, since there are no proofs of chaotic behavior.

There is, nevertheless, considerable theory for the case of ordinary differential equations (much less for partial equations) and several monographs are available expounding this theory. One of the best known is the book [6] by J. Guckenheimer and P. Holmes. There we learn again that the Soviets were ahead, since there is extensive discussion of the theories of V. K. Melnikov, from 1963 [11], and L. P. Shil'nikov, from 1968 [14]. The techniques of both of these pioneers were designed to reduce the study of certain systems of differential equations to the study of finite-dimensional maps. They show that imbedded in the phase space for these systems one can find a “horseshoe” map, which is a creation of S. Smale in the 1960s [15], and which enables one to show, for example, that the system in question has infinitely many periodic solutions. Not all workers accept this as a criterion for chaos, but it is the focus of much work, and in most cases, all that has been proved for smooth systems of differential equations.

The book by Guckenheimer and Holmes, like other literature in this field, is not easy reading. The theory was, and remains, incomplete and this is reflected in the unfinished nature of many of the results. For example, the following remark from their discussion of the concept of strange attractor is probably still valid.

“In trying to piece together a coherent picture of this situation, we enter a realm in which the theory remains in an unsatisfactory state. There are paradoxes in which different theorems appear to be steering us toward opposite conclusions.” . . .

We do not propose to resolve such problems here, or even to give an exposition of these fascinating concepts. Rather, we concentrate on the intuitive idea that a differential equation exhibits chaos if it has many solutions which are bounded and which are erratic and unpredictable in some sense. Within this limited scope there are a number of rigorous results, for smooth differential equations as well as for maps.

Generally these have been obtained by methods, like those of Melnikov and Shil'nikov, which in some way reduce the study of the differential equation to the study of a related map in a lower dimensional space. For example, if we are studying an autonomous system of three first-order ordinary differential equations, the related map is two-dimensional, taking some subset of the plane into itself. The goal of this paper is to show that results of this sort are accessible by different, and we think simpler, methods, which involve study of solutions of the differential equation directly rather than through a related map. We do this for a particular example, the equation for a pendulum with oscillating support, for illustration. The same technique can be applied to other equations, and some examples are given in [7] and [8].

II. EQUATION OF MOTION FOR A PENDULUM. First consider a simple pendulum, consisting of a mass m at the end of a massless rod of length l , which pivots on a frictionless support that forces the pendulum to move in a vertical plane (FIGURE 1).

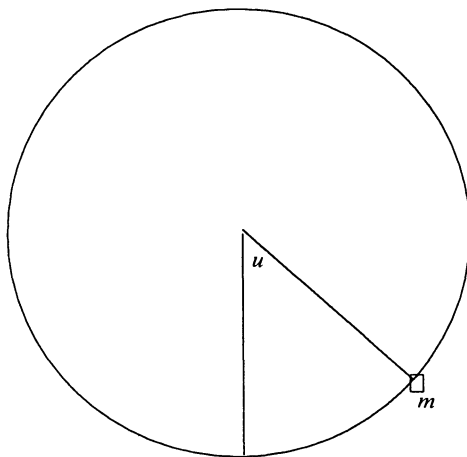


Figure 1. We consider a pendulum free to rotate in a full circle.

At time τ the rod makes an angle $u(\tau)$ with the vertical. The forces to be considered are the gravitational force mg and the damping due to air resistance as the pendulum swings. This is assumed to be proportional to the angular velocity ($du/d\tau = \dot{u}(\tau)$). Applying Newton's law of motion, we obtain the ordinary differential equation

$$ml\ddot{u} + c\dot{u} + mg \sin u = 0,$$

where c is the positive constant of proportionality. We immediately rescale to get the dimensionless version, by setting $t = \sqrt{g/l}\tau$ and $y(t) = u(\tau)$. Letting

$(dy/dt) = y'$, we get

$$y'' + ky' + \sin y = 0, \quad (1)$$

where $k = c/m\sqrt{gl}$.

The phase plane obtained by plotting $y'(t)$ against $y(t)$ is well known, and can be found in many introductory texts on ordinary differential equations, such as [2]. In FIGURE 2 we show the cases $k = 0$ and $0 < k < 2$.

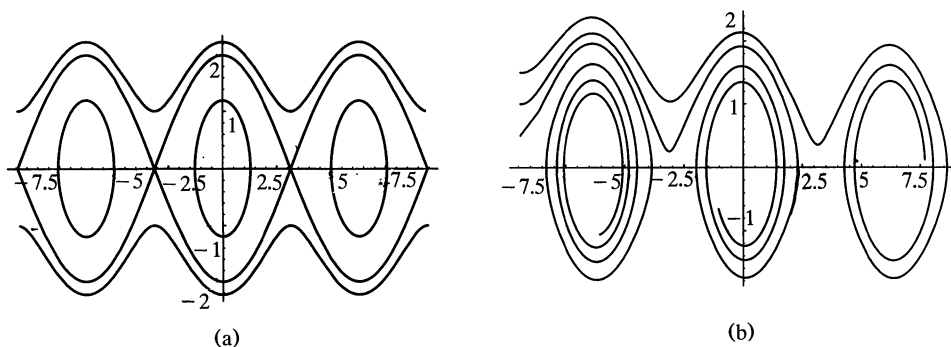


Figure 2. Two phase planes for the pendulum, one undamped, the other damped.

When $k = 0$ (no damping) the equation is called “conservative,” because there is a function of the solution which is conserved, or, in other words, is constant as t varies. This is the so-called “energy” function associated with the pendulum. If $(y(\cdot), y'(\cdot))$ is a solution, the associated energy is

$$E(t) = \frac{y'(t)^2}{2} - \cos y(t),$$

and is the usual energy function from physics, being a sum of the kinetic energy $y'(t)^2/2$ and a potential energy term $-\cos y(t)$, which is at a minimum when the pendulum is at its stable rest point, $y = 0$. To see that E is constant in the absence of damping, differentiate the expression for E and use (1) with $k = 0$.

FIGURE 2 illustrates conservation of energy when $k = 0$ because there is a family of orbits which are closed smooth curves, and represent periodic solutions. These are solutions with low energy. On the other hand, if the initial velocity is large, so that $E(t)$ is large, then the trajectory in phase space is unbounded, and represents a pendulum which continues to rotate in the same direction, making repeated complete rotations without loss of energy. It is important to observe that there are intermediate trajectories, such as the one which tends to the point $(-\pi, 0)$ in phase space as $t \rightarrow -\infty$, and to $(\pi, 0)$ as $t \rightarrow \infty$.

A trajectory in the phase plane of an autonomous differential equation represents many solutions, which can be characterized as passing through the same point in phase space but at different times. Corresponding to the trajectory connecting $(-\pi, 0)$ to $(\pi, 0)$ as t increases, which exists when $k = 0$, there is a unique solution y_0 of (1) such that

$$y_0(0) = 0, \quad \lim_{t \rightarrow \infty} y_0(t) = \pi, \quad \lim_{t \rightarrow -\infty} y_0(t) = -\pi. \quad (2)$$

Physically, this solution is approximated by a pendulum which starts from rest very close to the upright vertical position and makes almost a complete rotation, coming again close to the vertical position.

Chaotic motion is not possible in this model, whatever the value of k . To obtain more erratic solutions we must add some sort of forcing term. This can take various forms, but the one we study here results from assuming that the support of the pendulum is subject to a vertical motion, up and down, which is sinusoidal. This adds a force proportional to $\sin \epsilon t$ to the gravitational force. We make the assumption that the force on the support varies slowly with time, and also that the damping force is small. This results in the equation

$$y'' + \epsilon \delta y' + (1 + \gamma \sin \epsilon t) \sin y = 0 \quad (3)$$

where δ and γ are fixed positive numbers and ϵ is positive but small. To avoid the delicate case where the coefficient of $\sin y$ can be zero, we require that $0 < |\gamma| < 1$. Equation (3) was studied by S. Wiggins, in [16].

III. PREVIOUS RESULTS. To describe Wiggins' result, suppose that $y > 0$ represents displacement from rest in a counter-clockwise direction, and a full rotation occurs each time $y(t)$ crosses an odd multiple of π . He then shows that there is a $\Delta(\gamma) > 0$ such that the irregular behavior can occur if $0 \leq \delta < \Delta(\gamma)$ and ϵ is sufficiently small. We shall describe the function $\Delta(\cdot)$ shortly, but first let us specify the nature of this "irregular" behavior. Our measure of irregularity is that the pendulum makes a sequence of full rotations, alternating between clockwise and counter-clockwise rotations in an erratic manner. More precisely (without, however, yet specifying $\Delta(\gamma)$), we state this as a theorem.

Theorem 1 (Wiggins [16]). *Suppose that δ and γ are given numbers, with $0 < |\gamma| < 1$ and $0 \leq \delta < \Delta(\gamma)$. Then there is an $\epsilon_1 > 0$ such that for any ϵ with $0 < \epsilon < \epsilon_1$, and any finite or infinite sequence $\{m_j\}_{j=1,2,3,\dots}$ of positive integers, there is a solution of (3) such that the corresponding motion consists of exactly m_1 full clockwise rotations, followed by exactly m_2 full counter-clockwise rotations, m_3 full clockwise rotations, and so forth. If the sequence is finite, then eventually the pendulum stops making full rotations.*

It is common to refer to the solutions corresponding to infinite non-repeating sequences as "chaotic," though, as we said, some researchers prefer a stricter interpretation of this term.

Wiggins obtains this striking result by applying the technique of Melnikov to the equation (3). This requires extending Melnikov's original method, because previously the forcing term was required to be "small" in amplitude, whereas here the small parameter measures the frequency of the oscillation, not its amplitude. The necessary extension was also given by Palmer [12].

To define the function $\Delta(\cdot)$, Wiggins derives the appropriate "Melnikov" function for (3). While this concept has always seemed slightly mysterious to us, here it results from very standard energy methods involving the function $E(t)$. (See below.) Suppose that y_0 is the unique solution to (1) with $k = 0$ satisfying (2). Then

$$\Delta(\gamma) = |\gamma| \frac{\int_{-\infty}^{\infty} s y'_0(s) \sin y_0(s) ds}{\int_{-\infty}^{\infty} y'_0(s)^2 ds}.$$

In other words, chaotic solutions exist for sufficiently small $\epsilon > 0$ if

$$\int_{-\infty}^{\infty} I_{y_0}(s) ds > 0, \quad (4)$$

where

$$I_y(s) = -\delta y'(s)^2 + |\gamma|sy'(s)\sin y(s). \quad (5)$$

The left side of (4) is the Melnikov function for the equation (3). Since y_0 satisfies (1) with $k = 0$, we can set $\sin y_0(s) = -y_0''(s)$ in the formula for $\Delta(\gamma)$, integrate by parts, and use the boundary conditions to obtain that $\Delta(\gamma) = \frac{1}{2}|\gamma|$.

IV. PROOF WHEN $\delta = 0$. We will show how these results, and others, can be obtained by techniques which we feel are simpler than those used previously. Instead of studying Poincaré maps, we follow the solutions more directly, to determine how they vary as the initial conditions change. We need consider only initial conditions representing a pendulum which is released from a raised position, with zero initial velocity. The case $\delta = 0$ in (3) is particularly simple. Therefore we consider solutions to

$$y'' + (1 + \gamma \sin \epsilon t)\sin y = 0, \quad (6)$$

with initial conditions

$$y(0) = \alpha, \quad y'(0) = 0. \quad (7)$$

Sometimes we will denote the solution by y_α . The goal is to obtain complicated solutions by adjusting α . This is sometimes called a “shooting method,” because we attempt to “aim” the solution to get the desired behavior.

Shooting methods are topological, relying on separation theorems of some sort to distinguish between various types of behavior. As an example we prove a simple result about (6)–(7) which we will need later.

Lemma 1. *For sufficiently small $\epsilon > 0$ there is an $\hat{\alpha} \in (-\pi, 0)$ such that if $y = y_{\hat{\alpha}}$, then $y' > 0$ on $(0, \pi/2\epsilon]$ and $y(\pi/2\epsilon) = 0$. There is also an $\check{\alpha}$ such that $y' > 0$ on $(0, \pi/\epsilon]$ and $y(\pi/\epsilon) = 0$.*

Proof: We show how to get $\hat{\alpha}$; the argument for $\check{\alpha}$ is the same. If $-\pi < y < 0$, then $y'' > 0$, and so by choosing α in this range we ensure that $y' > 0$ as long as $y \leq 0$. Clearly, y crosses 0. Let

$$A = \left\{ \alpha \in (-\pi, 0) \mid y(t) = 0 \text{ before } t = \frac{\pi}{2\epsilon} \right\}$$

and

$$B = \left\{ \alpha \in (-\pi, 0) \mid y(t) = 0 \text{ after } t = \frac{\pi}{2\epsilon} \right\}.$$

Note that if $y(0) = -\pi$, $y'(0) = 0$, then y is constant, so that if α is very close to π , then y remains close to $-\pi$ for a long time before crossing 0. This shows that B is non-empty. To show that A is non-empty, consider a small negative α . As long as y is small, solutions of (6) are approximated by solutions of the linear equation $u'' + (1 + \gamma \sin \epsilon t)u = 0$. Solutions of this equation oscillate more quickly than solutions of $v'' + \sigma v = 0$ where $\sigma = 1 - |\gamma|$, and so cross zero in the interval $(0, \pi/\sqrt{\sigma})$. Hence if $2\epsilon < \sqrt{\sigma}$, then small negative α 's lie in B .

The crossings of zero are with $y' > 0$, and so the continuity of solutions with respect to α implies that A and B are open sets. They are obviously disjoint, and the connectedness of the interval $(-\pi, 0)$ implies that there is a point in this interval which is not in A or B . Such an α gives the solution $y_{\hat{\alpha}}$ described in the lemma.

Note that we make no assertion about whether $\hat{\alpha} > \check{\alpha}$ or vice versa, nor will we need any such result. Comparisons of this kind are generally difficult to obtain. They are used in uniqueness proofs, but this is a paper about existence.

A crucial fact about solutions of (3) is that if $\alpha = k\pi$ for some integer k , then the solution is constant. In fact a stronger statement is true:

(i) If, for some t_0 , $y(t_0) = k\pi$ and $y'(t_0) = 0$, then $y(t) = k\pi$ for all t .

This follows from the uniqueness theorem for initial value problems for ordinary differential equations. The solution $y \equiv k\pi$ is the unique solution satisfying the conditions $y = k\pi$, $y' = 0$ at the point $t = t_0$.

One reason that the case $\delta = 0$ is particularly simple is that in this case some solutions have certain symmetries, around points T_n/ϵ , where n is an integer and $T_n = (2n + 1)\pi/2$. These are as follows.

(ii) If $y(T_n/\epsilon) = k\pi$ for some integer k , then

$$y\left(\frac{T_n}{\epsilon} + s\right) - k\pi = k\pi - y\left(\frac{T_n}{\epsilon} - s\right),$$

for all s .

(iii) If $y'(T_n/\epsilon) = 0$, then

$$y\left(\frac{T_n}{\epsilon} + s\right) = y\left(\frac{T_n}{\epsilon} - s\right)$$

for all s .

These are also proved by using the uniqueness of solutions to initial value problems. Note that the solution $y_{\hat{\alpha}}$ found in Lemma 1 is antisymmetric around $\pi/2\epsilon$, and therefore it increases up to π/ϵ , where it has a maximum in the region $0 < y < \pi$ and then starts to decrease.

The only detailed analysis required to prove Theorem 1 when $\delta = 0$ is used to obtain the following lemma.

Lemma 2. *For sufficiently small $\epsilon > 0$, there is some $\underline{\alpha}$ with $-\pi < \underline{\alpha} < 0$, such that $y_{\underline{\alpha}}$ increases monotonically on some interval $[0, t_0]$ and $y_{\underline{\alpha}}(t_0) = \pi$.*

The proof of Lemma 2 is quite simple in the situation we are considering. We give an outline below. Here we want to show how it is used in conjunction with a shooting technique to prove Theorem 1.

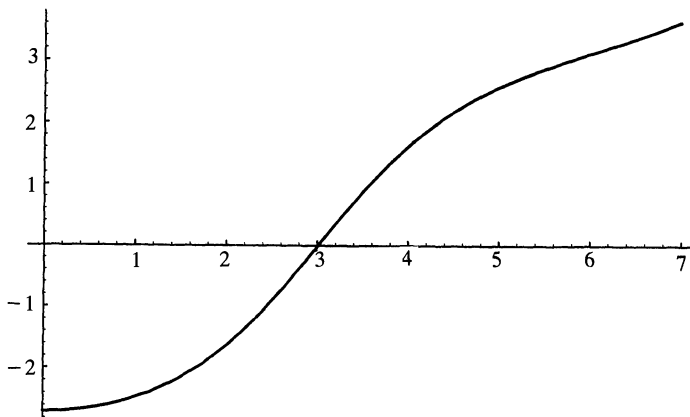


Figure 3. A graph of a solution behaving as described in Lemma 2.

Suppose that the first complete rotation of the pendulum is to be in the direction of positive y ; i.e. we want the solution to cross $y = \pi$ before any possible crossing of $y = -\pi$. We begin by choosing $\alpha = \underline{\alpha}$ as in Lemma 2. The corresponding solution crosses $y = \pi$ at a point t_0 , with $y' > 0$, and since $y'' > 0$ when $\pi < y < 2\pi$, there must be a $t_1 > t_0$ such that $y' > 0$ on $(0, t_1]$, and $y(t_1) = 2\pi$. Since $y'(t_1) > 0$, the implicit function theorem implies that there is a smooth function $t_1(\alpha)$, defined in a neighborhood of $\underline{\alpha}$ by the equation $y_\alpha(t_1(\alpha)) = 2\pi$.

However, recalling the properties of y in Lemma 1 when $\alpha = \hat{\alpha}$, we see that $t_1(\hat{\alpha})$ is not defined. Suppose for convenience that $\hat{\alpha} > \underline{\alpha}$. Then the function $t_1(\cdot)$ is continuous in some maximal interval of the form $[\underline{\alpha}, \bar{\alpha})$, where $\bar{\alpha} \leq \hat{\alpha} < 0$. Since $t_1(\cdot)$ can be extended continuously to an open neighborhood of any point where it is defined, we conclude that

$$\lim_{\alpha \rightarrow \bar{\alpha}^-} t_1(\alpha) = \infty. \quad (8)$$

This result depends on property (i) above, for otherwise a crossing of 2π might disappear by the solution becoming tangent to this line for some α .

To illustrate the idea of the proof, suppose that $m_1 = 2$, so that we want y to cross 3π before recrossing π . This is accomplished by moving α from $\underline{\alpha}$ towards $\bar{\alpha}$. Since $t_1(\cdot)$ is continuous, it follows from (8) that there is an $\alpha_1 \in (\underline{\alpha}, \bar{\alpha})$ such that $t_1(\alpha_1) = T_{n_1}/\epsilon$, for some integer n_1 . We now apply the anti-symmetry principle (ii). Since $y'(0) = 0$, $-\pi < y(0) < 0$, and y increases monotonically to reach 2π at T_{n_1}/ϵ , it must continue to increase monotonically until it crosses 3π and 4π , after which it has a maximum before any possible crossing of 5π . We have then accomplished the first step of achieving exactly two counter-clockwise rotations, and we next wish to obtain a clockwise rotation, since we are assuming that $m_2 \geq 1$.

Let $t_2(\alpha_1)$ be the first $t > 0$ where y_{α_1} has a local maximum. As we have noted, $4\pi < y_{\alpha_1}(t_2(\alpha_1)) < 5\pi$. Then $t_2(\cdot)$ can be extended as a continuous function in some neighborhood of α_1 , as a solution of the equation $y'_\alpha(t) = 0$. As long as $t_2(\cdot)$ is continuous, $y_\alpha(t_2(\alpha))$ must lie in the interval $(4\pi, 5\pi)$, since (i) implies that the maximum cannot leave this interval by means of a tangency. Moreover, $t_2(\cdot)$ can be extended continuously to a maximum interval of the form $[\alpha_1, \bar{\alpha}_1)$, where $\alpha_1 < \bar{\alpha}_1 \leq \bar{\alpha}$, and

$$\lim_{\alpha \rightarrow \bar{\alpha}_1^-} t_2(\alpha) = \infty.$$

Therefore, we can find $\alpha_2 \in (\alpha_1, \bar{\alpha}_1)$ such that $t_2(\alpha_2) = T_{n_2}/\epsilon$, for some integer n_2 . It is important to note that $t_1(\alpha_2)$ is not of the form T_n/ϵ , but it is still defined, as the first point where y_{α_2} crosses 2π . By (iii), y_{α_2} is symmetric around $t_2(\alpha_2)$, and so it must descend from its maximum there to recross 3π and π . If $m_2 = 2$ then this completes the second step of the induction process, since with the choice we have made of α_2 , the next extremum of y_{α_2} is a minimum at $t = 2t_2(\alpha_2)$, where $y = y(0) \in (-\pi, 0)$. If $m_2 \neq 2$, we must adjust α further. As we adjust α , the various crossing points defined so far will change. However at each adjustment we only move α enough to bring the most recently defined crossing point or critical point to one of the points of symmetry or antisymmetry on the t axis. All the earlier such points remain bounded, and hence continuous in α , and none of the critical points can cross any line $y = k\pi$.

The case $m_2 = 5$ is typical. So far we have a solution which increases from its starting point in the interval $(-\pi, 0)$ past π and 3π , and then decreases back past 0. We can increase α still further so that this (downward) crossing of 0 is at one of

the points T_n/ϵ . Then the antisymmetry property (ii) shows that the solution must continue to decrease past $-\pi$ and -3π , but not past -5π .

This means the pendulum makes four counter-clockwise rotations, and we want exactly five. This is achieved by a further increase of α so that the point of antisymmetry is where the solution crosses $-\pi$. Since this is half-way between 3π and -5π , and the solution has a maximum in the earlier interval where it was between 4π and 5π , it will now have a minimum between -6π and -7π . It crosses -5π , but not -7π , which completes the second step.

Continuing this process, we obtain an increasing sequence of α 's which is bounded above by $\bar{\alpha}$. This sequence must have a limit lying in the interval $(-\pi, 0)$, and this limiting value of α gives the solution we were after. Note that at each step, when we adjust α so that some crossing or extremal point lies at some odd multiple of $\pi/2\epsilon$, there are infinitely many choices of α , since there are infinitely many such odd multiples. This gives, for each sequence $\{m_i\}$, an infinite number of solutions of the desired type.

V. CHAOS AT PARTICULAR PARAMETER VALUES. It is apparent from the proof of Theorem 1 that, for $\delta = 0$, chaotic solutions exist if there is a solution such that $-\pi < y(0) < 0$, $y'(0) = 0$, and $y(T) = \pi$ for some $T > 0$. For particular parameter values this can be verified by following only one trajectory for a finite time interval. Therefore, in principle, rigorous estimates can be made which allow a proof that chaotic solutions exist for precise values of the parameters, rather than "for sufficiently small ϵ ," as in Theorem 1. First we have to locate a good candidate for the parameter values. Standard numerical experimentation can easily be done on a personal computer, for example with the software Phsplan [4]. It is quickly determined that for $\epsilon = 0.1$, $\gamma = -0.5$, the solution with $y(0) = -2.7$, $y'(0) = 0$ increases monotonically until it crosses π , at approximately $t = 6.1$.

We then make use of a technique in numerical analysis called "interval arithmetic." In interval arithmetic, the computer is programmed to include error estimates in all arithmetic computations. An exposition is given in [1]. In addition to roundoff error it is necessary to allow for the truncation error introduced by the numerical method. Programs can be written to do this. We used PBASIC [1], which uses precise interval arithmetic, to do a completely rigorous integration of the equation with the parameter values and initial conditions found approximately using standard floating point computations. The result confirmed that the solution does indeed cross π , and therefore that with these parameter values there are solutions of arbitrary complexity, as described in the theorem.

VI. EXTENSIONS. The outline above of the proof of Theorem 1 makes it appear that symmetry is crucial for this result. But this is misleading. Symmetry considerations shorten the proof, but the heart of our method is the topological shooting principle. This is fortunate, for as soon as we add a damping term, as in (3), properties (ii) and (iii) do not hold. Shooting works because (i) is still valid. We need a slight extension of Lemma 2, but this is not difficult.

Physical intuition may cause doubts about this, because damping reduces energy and tends to stop the complete rotations. However the oscillation of the support can add energy if it is timed correctly. It is rather like pushing a child's swing. If the pushes occur in the direction of motion, the swing will go higher, despite air resistance. If the resisting forces are not too high we can indeed push a swing over the top, (though perhaps the child will be better off if we do not). Do not carry this analogy too far, however. A very important difference is that in our case the

motion of the support is determined ahead of time, and not adjusted to fit the motion of the pendulum.

In fact, even the periodicity of the forcing is not required. We can consider more general equations

$$y'' + \epsilon \delta y' + p(\epsilon t) \sin y = 0 \quad (9)$$

where the positive smooth function p increases and decreases in some fashion. For example, here is a set of sufficient conditions on p .

(a) p , $1/p$, p' and p'' exist and are bounded on $[0, \infty)$.

(b) There are sequences $\{t_j\}$ and $\{\tau_j\}$ tending to infinity and a $c > 0$ such that $p'(t_j) \geq c$ and $p'(\tau_j) \leq -c$.

This includes almost periodic functions and many others, and goes beyond what has been found using Poincaré maps, for it is difficult to define such a map usefully when the equation depends explicitly on time in an irregular fashion.

The damping term $\epsilon \delta y'$ in (9) is responsible for the introduction of the Melnikov function as described earlier. This enters into the proof of Lemma 2 when $\delta > 0$, but plays no role in the shooting part of the argument. We conclude this paper with a brief discussion of how to prove Lemma 2.

VII. OUTLINE OF PROOF OF LEMMA 2. The basic idea is to consider the energy $E(t)$, as defined earlier. The damping term tends to reduce E (makes $E' < 0$ when $y' \neq 0$), while the oscillation in the support may increase or decrease E , depending on the relative direction of movement of the support and the pendulum at any given time. Suppose again that $\delta = 0$ and $p(s) = 1 + \gamma \sin s$. We find using (6) that $E'(t) = -\gamma y'(t) \sin \epsilon t \sin y(t)$. To prove Lemma 2, we show that the initial position α can be adjusted so that $E(t)$ increases enough during the first swing to get the pendulum “over the top” for one complete revolution.

Suppose that $\gamma > 0$. Then we can choose $\alpha = \check{\alpha}$, which was found in Lemma 1. In this case we have $\sin \epsilon t > 0$ and $\sin y < 0$ in the interval $(0, \pi/\epsilon)$, and both of these quantities change sign at this point. It follows that E increases on the entire interval $(0, 2\pi/\epsilon)$, as long as $y' > 0$, $y < \pi$.

We can estimate the change in E while $y' > 0$, $y < \pi$ as follows. If $\pi/\epsilon \leq t \leq 2\pi/\epsilon$, then

$$\begin{aligned} E(t) &= E(0) + \int_0^t \{-\gamma y'(s) \sin \epsilon s \sin y(s)\} ds \\ &> E(0) + \int_{\pi/\epsilon-1}^{\pi/\epsilon} \{-\gamma y'(s) \sin \epsilon s \sin y(s)\} ds. \end{aligned} \quad (10)$$

The second term on the right is estimated by proving a simple lemma showing that as $\epsilon \rightarrow 0$, the solution y found in Lemma 1 must tend to $y_0(t - \pi/\epsilon)$, where y_0 is the unique solution of (1)–(2) with $k = 0$. That is, for example,

$$\max_{\pi/\epsilon-1 \leq t \leq \pi/\epsilon} \left| y(t) - y_0\left(t - \frac{\pi}{\epsilon}\right) \right| \rightarrow 0.$$

Setting $s - \pi/\epsilon = \sigma$, and noting that $\sin \epsilon \sigma \rightarrow \epsilon \sigma$ uniformly on compact σ -intervals, we find that the second term on the right of (10) is asymptotically like $\int_{-1}^0 \gamma \epsilon \sigma y'_0(\sigma) \sin y_0(\sigma) d\sigma = \mu \epsilon$ where μ is a positive number. We also need an easy estimate for $E(0) = -\cos \check{\alpha}$. Without going into further details, this enables us to obtain quickly an estimate of the form $y'(t) \geq k\sqrt{\epsilon}$ for $t \geq \pi/\epsilon$, implying that y rises above π before $t = 2\pi/\epsilon$.

When δ is positive, the rate of change of energy is given by $E'(t) = -\delta \epsilon y'^2 - \gamma y' \sin \epsilon t \sin y$. The first term causes E to decrease, while the sign of the second

term alternates. The analysis of the net change of E in the same situation as above, where $y = 0$ when $t = \pi/\epsilon$, is a little more complicated and leads to a consideration of the Melnikov function. The case of a non-periodic forcing is no harder; periodicity is irrelevant to Lemma 2.

VIII. FINAL REMARKS. Theorem 1 and the extensions described in the last section by no means tell the whole story, even for (6). While we show that there is an uncountable number of erratic solutions, our numerical simulations indicate that most solutions eventually settle down into small oscillations around an even multiple of π , which represents the downward vertical position of the pendulum. We do not know of a proof of this, however.

Also, we do not know whether the results can be extended to include larger values of ϵ . Our standard numerical computations (not rigorous) show that the crucial solution of Lemma 2 exists at least out to $\epsilon = 50$. For very large ϵ it seems that another subtle effect enters in, the so-called “exponential splitting of separatrices” [5]. This is certainly beyond the scope of this paper. What we would like to show is that the phenomenon occurs for all values of ϵ , and this seems to require some new estimates.

REFERENCES

1. O. Aberth, *Precise Numerical Analysis*, William C. Brown Publishers, Dubuque, Iowa, 1988.
2. F. Brauer and J. Nohel, *Introduction to Differential Equations with Applications*, Harper and Row, 1985.
3. R. L. Devaney, *Chaotic Dynamical Systems*, Addison Wesley, 1989, 2nd edition.
4. G. B. Ermentrout, *PhasePlane, the Dynamical Systems Tool*, Brooks/Cole Publishing, Pacific Grove, CA, 1990.
5. E. Fontich and C. Simó, The Splitting of Separatrices for Analytic Diffeomorphisms, *Ergodic Theory and Dynamical Systems* 10 (1990), 295–318.
6. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag (NY), 1982.
7. S. Hastings and J. B. McLeod, On the Periodic Solutions of a Forced Second-Order Equation, to appear in *Nonlinear Science*.
8. S. Hastings and W. Troy, Oscillating Solutions of the Falkner-Skan Equation for Positive β , *J. Diff. Eqns.* 71 (1988), 123–144.
9. Y. Li and J. Yorke, Period Three Implies Chaos, *American Mathematical Monthly* 82 (1975), 985–992.
10. E. N. Lorenz, Deterministic Non-periodic Flow, *J. Atmospheric Sci.* 20 (1963), 130–141.
11. V. K. Melnikov, On the Stability of the Center for Time Periodic Solutions, *Trans. Moscow Math. Soc.* (Trudy), 12 (1963), 3–52.
12. K. J. Palmer, Transversal Heteroclinic Points and Cherry’s Example of a Nonintegrable Hamiltonian System, *J. Diff. Eqns* 65 (1986), 321–360.
13. A. N. Sharkovsky, Coexistence of Cycles of a Continuous Map of a Line into itself, *Ukrainian Math. J.* 16 (1964), 61–71.
14. L. P. Shil’nikov, On the Generation of a Periodic Motion from Trajectories Doubly Asymptotic to an Equilibrium State of Saddle Type, *Math. USSR-Sb.* 6 (1968), 427–438.
15. S. Smale, Diffeomorphisms with Many Periodic Points, *Differential and Combinatorial Topology*, Princeton Univ. Press, Princeton, NJ, 1965, 63–80.
16. S. Wiggins, On the Detection and Dynamical Consequences of Orbits Homoclinic to Hyperbolic Periodic Orbits and Normally Hyperbolic Invariant Tori in a Class of Ordinary Differential Equations, *SIAM J. Appl. Math.* 48 (1988), 262–285.

Department of Mathematics & Statistics
University of Pittsburgh
Pittsburgh, PA 15260
sph@mthsn4.math.pitt.edu

An Application for the Curiosity $(\log_x N)'$

David A. Wagstaff, Theodore A. Norman,
and Douglas M. Campbell

Students often believe that differentiation formulas such as

$$(*) \quad [\log_x N]' = -(\log_x N)/(x \ln x)$$

are mere curiosities. We present a practical application of (*).

In practice, unsorted data files on a hard disk may be extremely large (e.g. 40 megabytes), while available RAM (Random Access Memory) on many personal computers is small (e.g. 1 megabyte). There is a simple strategy to sort such a file:

1. Divide it into chunks which are the size of RAM (for our example, 1 megabyte chunks). For each chunk, read the chunk into RAM, sort it by one's favorite internal sort, and write the chunk back to the hard disk as a separate file (see [1], p. 263–p. 270; [2] chapter 5.4, Theorem L, page 371).

2. Then, as FIGURE 1 indicates, groups of x of these sorted chunks are merged into a larger sorted chunk, and written back to the hard disk. This continues until the final merge, in which only x huge chunks remain and they are merged into the final sorted file.

Although this example is a considerable simplification of the real problem, the most time consuming part of this operation is the *seek*, in which the access mechanism of the hard disk is moved to the proper track on the hard disk to read the information. The question arises, what value of x will minimize the number of seeks? Furthermore, how does the value of x depend on the file size and the available RAM?

Theorem. *Let a file have N records and let the computer's RAM hold M records. The value of x to minimize the number of seeks in an x -way merge is 3, independent of N and M .*

Proof: Sub-divide the N records into $R = N/M$ files (chunks), which are read into memory, internally sorted, and written to the hard disk to form the top row of Figure 1. (By adding specially marked dummy records, we may assume that R is a power of x .) The system of x -way merges of Figure 1 has $\log_x R$ levels. At each level, the contents of the original file have to be read into x buffers. Since RAM can only hold M records, each of the x buffers holds M/x records. Each time a buffer is filled, a seek is required. Therefore, the number of seeks per level is $N/(M/x)$ which is xR . The total number of seeks, y , is the number of seeks per level times the number of levels:

$$y = xR \log_x R.$$

Taking the derivative with respect to x we see that $y' = R \log_x R [-1 + \ln x]/\ln x$, which is zero when $x = e$. Since y' goes from negative to positive at $x = e$, the function y has its minimum at $x = e$, independent of N and M . But the

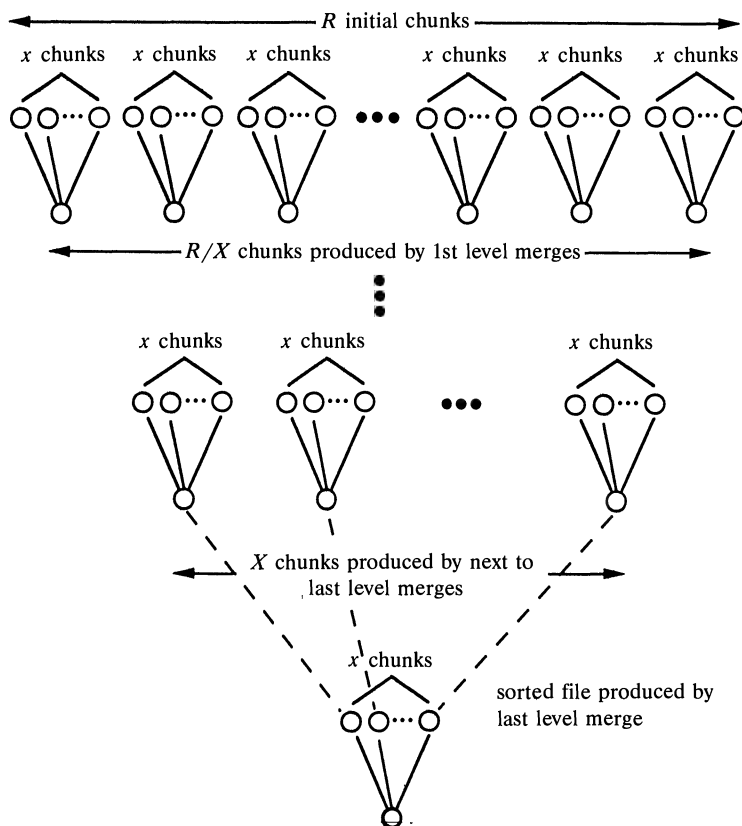


Figure 1

value of x in an x -way merge must be an integer. Writing R as 2^z , for some $z > 0$, we note that

$$\begin{aligned}
 y(2) - y(3) &= 2R \log_2 R - 3R \log_3 R \\
 &= R(2z - 3z \log_3 2) \\
 &= Rz(\log_3 9 - \log_3 8),
 \end{aligned}$$

which confirms that a 3-way merge minimizes the number of seeks.

To derive (*), we note that $y = \log_x N$ can be rewritten as $x^y = N$ and thereby as $y \ln x = \ln N$. Taking the derivative yields $y' \ln x + y/x = 0$, which when solved for y' yields (*).

REFERENCES

1. Michael J. Folk, Bill Zoelick, "File Structures: A Conceptual Toolkit", Addison-Wesley, Reading, Mass, 1987.
2. Donald E. Knuth, "The Art Of Computer Programming, Volume 3, Sorting and Searching", Addison-Wesley, Reading, Mass, 1973.

Computer Science Department
Brigham Young University
Provo, UT 84602
campbell@cs.byu.edu

Vandermonde Strikes Again

Miriam Schapiro Groszof and Geraldine Taiani

The search for “cute” proofs—those in which arguments or techniques that at first glance appear completely unrelated in substance are used to establish familiar results—provides healthful exercise for both students and experts. One of the more versatile tools for this purpose is the Vandermonde matrix and its determinant. This mathematical object has an honorable history dating from the late 18th century [1] and has enjoyed sporadic revivals of interest [3], [5]. In the past, most undergraduate major courses included its applications in the proof that values at $n + 1$ distinct points uniquely determine a monic polynomial of degree n [7] and in the definition of the signature of a permutation [4]. We present here a novel, indeed unexpected, application of the Vandermonde.

A theorem of Abel [2] states: *if $P(x), Q(x)$ are any two polynomials such that $\deg Q = n \geq 3$, Q has no multiple roots, and $\deg P = m \leq n - 2$, then*

$$(A) \sum \frac{P(r_i)}{Q'(r_i)} = 0$$

where the summation is over all n distinct roots r_i of Q . (As usual, Q' denotes the derivative of Q .)

Abel’s original proof (of a more complicated result) uses integrals. The modern standard proof is based on residue theory. Since $\deg P \leq \deg Q - 2$,

$$\left| \int_{|z|=R} \frac{P(z)}{Q(z)} dz \right| \leq \frac{C}{R} \quad \text{for some constant } C > 0 \text{ and } R \text{ sufficiently large;}$$

hence

$$\lim_{R \rightarrow \infty} \int_{|z|=R} \frac{P(z)}{Q(z)} dz = 0.$$

However,

$$\int_{|z|=R} \frac{P(z)}{Q(z)} dz = 2\pi i \cdot \left(\text{sum of residues of } \frac{P(z)}{Q(z)} \text{ inside } |z| < R \right),$$

and since $P(z)/Q(z)$ has simple poles at the roots of Q the residues of $P(z)/Q(z)$ inside $|z| = R$ are precisely

$$\left\{ \frac{P(r_i)}{Q'(r_i)}, r_i \text{ roots of } Q \text{ in } |z| = R \right\}.$$

The desired (A) follows.

We have found an algebraic proof of (A) in the spirit of classical theory of equations. Strictly speaking it requires no complex analysis. Given polynomial

Q with $\deg Q = n \geq 3$ and distinct (real or complex) roots r_1, r_2, \dots, r_n , assume w.l.o.g. Q is monic so $Q(x) = (x - r_1)(x - r_2) \cdots (x - r_n)$. Then $Q'(x) = \sum_i (x - r_1)(x - r_2) \cdots \widehat{(x - r_i)} \cdots (x - r_n)$, that is, each summand has one factor, $(x - r_i)$, omitted. Thus,

$$\begin{aligned} Q'(r_i) &= (r_i - r_1)(r_i - r_2) \cdots (r_i - r_{i-1})(r_i - r_{i+1}) \cdots (r_i - r_n) \\ &= (-1)^{n-i} (r_i - r_1)(r_i - r_2) \cdots (r_i - r_{i-1})(r_{i+1} - r_i) \cdots (r_n - r_i) \\ &= (-1)^{n-i} \frac{\prod_{j>k} (r_j - r_k)}{\prod_{\substack{j>k \\ j, k \neq i}} (r_j - r_k)}. \end{aligned}$$

Recall now the Vandermonde determinant [6]

$$V_n(a_1, \dots, a_n) = \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ a_1^{n-2} & a_2^{n-2} & \cdots & a_n^{n-2} \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix}$$

is a polynomial of degree $n(n-1)/2$ in the n variables a_1, \dots, a_n ; it can be written as the product $(a_2 - a_1) \cdots (a_n - a_{n-1}) = \prod_{j>k} (a_j - a_k)$. $V_n(a_1, \dots, a_n) = 0$ if and only if $a_k = a_j$ for some $k \neq j$. Moreover, the minors of entries in row n are themselves Vandermondes: in particular, the minor of a_i^{n-1} is $V_{n-1}(a_1, \dots, \hat{a}_i, \dots, a_n)$ which is precisely $\prod_{j>k, j, k \neq i} (a_j - a_k)$. Hence,

$$Q'(r_i) = (-1)^{n-i} \frac{V_n(r_1, \dots, r_n)}{V_{n-1}(r_1, \dots, \hat{r}_i, \dots, r_n)}.$$

Now given polynomial P with $0 \leq m = \deg P \leq n-2$

$$\begin{aligned} \sum_i \frac{P(r_i)}{Q'(r_i)} &= \sum_i (-1)^{n-i} P(r_i) \frac{V_{n-1}(r_1, \dots, \hat{r}_i, \dots, r_n)}{V_n(r_1, \dots, r_n)} \\ &= \frac{(-1)^{n-1}}{V_n(r_1, \dots, r_n)} \sum_i (-1)^{i-1} P(r_i) V_{n-1}(r_1, \dots, \hat{r}_i, \dots, r_n) \\ &= \frac{(-1)^{n-1}}{V_n(r_1, \dots, r_n)} \cdot (-1)^{n-1} \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1 & r_2 & \cdots & r_n \\ \vdots & \vdots & \cdots & \vdots \\ r_1^{n-2} & r_2^{n-2} & \cdots & r_n^{n-2} \\ P(r_1) & P(r_2) & \cdots & P(r_n) \end{bmatrix}. \end{aligned}$$

However, $P(x)$ has degree $\leq n-2$ so that each $P(r_i)$ is the same linear combination of $1, r_i, r_i^2, \dots, r_i^{n-2}$ and hence the determinant is zero, as desired.

Note that this result (and Abel's proof but not ours) is true when $n = 2, m = 0$; $n = 2, P \equiv 0$; or $n = 1, P \equiv 0$. These cases are easily proved by differentiation and substitution in (A).

We have noticed that recent texts ignore the Vandermonde so that even our advanced students have never heard of it, nor its discoverer, nor indeed the entire category of problems which drove the work of Lagrange, Galois, Abel and their heirs. The topic (along with the rest of theory of equations, now “lost”) is well suited to independent study, a mini-course or a special project, if there is no room for it in the modern over-crowded pregraduate major sequence.

We wish to thank the referee for a useful comment.

REFERENCES

1. Florian Cajori, *A History of Mathematics*, (ed. 2), Macmillan, NY, 1919, p. 266.
2. P. Griffiths, Variations on a theorem of Abel, *Inventiones* 35 (1976), 321–390.
3. Allen Klinger, The Vandermonde matrix, *Am. Math. Monthly* 74 (1967), 571–574.
4. Lowell J. Paige, and J. Dean Swift, *Elements of Linear Algebra*, Ginn & Co., Boston, 1961, p. 109.
5. Joseph J. Rushanan, On the Vandermonde matrix, *Am. Math. Monthly* 96 (1989), 921–924.
6. J. V. Uspensky, *Theory of Equations*, McGraw-Hill, NY, 1938, p. 214.
7. Louis Weisner, *Introduction to the Theory of Equations*, Macmillan, NY, 1938, p. 56.

Department of Mathematics
Yeshiva University
New York, NY 10016

Department of Mathematics
Pace University
New York, NY 10038

More on pi

I may well have made a mistake in my note *How to Make Pi Equal to Three* in the February 1992 issue of this *Monthly*, but it was not the mistake that Professor Dario Castellanos points out in his recent letter (*Monthly*, January 1993), because I did not take my ruler aboard the spinning circle.

Professor Castellanos takes a ruler on board a spinning circle, and uses it to measure the circumference of a stationary circle. Naturally, he gets a value of π greater than usual.

I used a stationary ruler to measure the circumference of a spinning circle, so I got a value of π less than usual.

The trouble is I used special relativity instead of general relativity. A spinning circle is an accelerated system, which calls for general relativity. I assumed that for large radius we could approximate circular motion with straight line motion, and proceeded accordingly.

But I never took my ruler on board the spinning circle.

Along these same lines, here is a paradox I cannot explain. Suppose a train on a circular track is so long that reaches all the way around, and the caboose is hitched to the locomotive. If the train travels near the speed of light, each car decreases in length, so we have a short train filling a long track. What happens?

—Rick Norwood
Department of Mathematics
East Tennessee State University
Johnson City TN 37614

NOTES

Edited by: John Duncan

Embedding Countable Groups in 2-Generator Groups

Fred Galvin

The aim of this note is to popularize a simple proof, due to Neumann and Neumann [5], of the fact that every countable group is embeddable in a 2-generator group. This was first proved by Higman, Neumann, and Neumann [2, Theorem IV] and (independently) Freudenthal [2, p. 254], using free products with amalgamations. The proof given by Neumann and Neumann [5] used wreath products, which are widely regarded as no less terrifying than free products with amalgamations. I am indebted to Professors A. M. W. Glass, G. Higman, and P. M. Neumann, each of whom pointed out (in response to an earlier version of this note) that the Neumann-Neumann proof is really quite simple, and that it can easily be expressed directly in terms of permutations. Here, then, is a short proof that assumes no more background than is needed to understand the statement of the theorem.

As usual, \mathbb{Z} is the set of integers and \mathbb{N} is the set of natural numbers; $\langle a, b \rangle$ is the group generated by a and b ; $\text{Sym}(\Omega)$ is the group of all permutations of a set Ω ; permutations are regarded as right operators, and are composed from left to right.

Theorem 1. *Every countable group is embeddable in a 2-generator group.*

Proof: Consider a countable group $G = \{g_1, g_3, g_5, \dots\}$; the elements are indexed by odd positive integers. We may assume that G is a subgroup of $\text{Sym}(\mathbb{N})$. Define permutations a and b in $\text{Sym}(\mathbb{Z} \times \mathbb{Z} \times \mathbb{N})$ by setting $(m, n, p)a = (m + 1, n, p)$ and

$$(m, n, p)b = \begin{cases} (m, n + 1, p) & \text{if } m = 0; \\ (m, n, pg_m) & \text{if } m \text{ is odd, } m > 0, n \geq 0; \\ (m, n, p) & \text{otherwise.} \end{cases}$$

Let $b_i = a^i b a^{-i}$ and $\hat{g}_i = b_i b^{-1} b_i^{-1} b$ for $i = 1, 3, 5, \dots$. Straightforward (if slightly tedious) calculation shows that $(m, n, p)\hat{g}_i = (0, 0, pg_i)$ if $m = n = 0$, while $(m, n, p)\hat{g}_i = (m, n, p)$ otherwise. Thus $\hat{G} = \{\hat{g}_i: i = 1, 3, 5, \dots\}$ is a subgroup of $\langle a, b \rangle$ isomorphic to G .

It may have occurred to the reader to wonder whether Theorem 1 can be proved by just showing that any countable subgroup of a symmetric group $S = \text{Sym}(\Omega)$ is contained in a 2-generator subgroup of S . In fact, this was a question of Wagon [8]. An old theorem of Sierpiński [7, 9] says that any countable set of

selfmaps of an infinite set Ω is contained in the semigroup generated by two selfmaps of Ω ; Wagon asked whether one could replace “selfmaps” by “permutations” in Sierpiński’s theorem. The answer is yes [1], but the proof is a bit more involved and will appear elsewhere.

The two generators in the proof of Theorem 1 were both of infinite order. B. H. Neumann [4, p. 541] remarked that every countable group is embeddable in a 2-generator group with generators of prescribed orders $q \geq 8$ and $r \geq 2$; this was improved by Levin [3] to $q \geq 3$ and $r \geq 2$, which is the best possible result in this direction. The proof of Levin’s result is a little too complicated to give here; however, we can get two generators of finite order by modifying the proof of Theorem 1. We need the following easy lemma:

Lemma 1 [6, Exercise 10.1.17, p. 259]. *Every permutation is the product of two involutions.*

Proof: It suffices to consider the case of a permutation consisting of a single (finite or infinite) cycle. Note, e.g., that a 6-cycle is obtained by multiplying the involutions $(1, 2)(3, 4)(5, 6)$ and $(2, 3)(4, 5)$. This example can easily be generalized to get cycles of any desired length.

Theorem 2. *Every countable group is embeddable in a 2-generator group with one generator of order 11 and the other of order 2.*

Proof: Let G be a countable group. We may assume that G is a subgroup of $\text{Sym}(\mathbb{N})$; moreover, by Theorem 1 and Lemma 1, we may assume that G is generated by four involutions, which we call g_3, g_5, g_7 , and g_9 . Define permutations a and b in $\text{Sym}(\mathbb{Z}_{11} \times \mathbb{Z} \times \mathbb{N})$, of orders 11 and 2 respectively, by setting $(m, n, p)a = (m + 1, n, p)$ and

$$(m, n, p)b = \begin{cases} (m, n + (-1)^n, p) & \text{if } m = 0; \\ (m, n - (-1)^n, p) & \text{if } m = 1; \\ (m, n, pg_m) & \text{if } m \in \{3, 5, 7, 9\}, n \equiv 0 \pmod{4}, n \geq 0; \\ (m, n, p) & \text{otherwise.} \end{cases}$$

Let $c = (baba^{-1})^2$. Note that $(0, n, p)c = (0, n + 4(-1)^n, p)$, while $(m, n, p)c = (m, n, p)$ if $m \neq 0$. Let $b_i = a^i b a^{-i}$ and $\hat{g}_i = b_i c^{-1} b_i c$ for $i = 3, 5, 7, 9$. Then $(m, n, p)\hat{g}_i = (0, 0, pg_i)$ if $m = n = 0$, while $(m, n, p)\hat{g}_i = (m, n, p)$ otherwise. Hence $\langle \hat{g}_3, \hat{g}_5, \hat{g}_7, \hat{g}_9 \rangle$ is a subgroup of $\langle a, b \rangle$ isomorphic to G .

Theorem 3. *Every countable group is embeddable in a 2-generator group with one generator of prescribed order $q \geq 5$ and the other of order 2.*

Proof: By Theorem 2 and Lemma 1, we may assume that the given countable group is a subgroup of $\text{Sym}(\mathbb{N})$ generated by three involutions, which we call g_2, g_3 , and g_4 . Define permutations a and b in $\text{Sym}(\mathbb{Z}_q \times \mathbb{Z} \times \mathbb{N})$, of orders q and 2

respectively, by setting $(m, n, p)a = (m + 1, n, p)$ and

$$(m, n, p)b = \begin{cases} (m, n + (-1)^n, p) & \text{if } m = 0; \\ (m, n - (-1)^n, p) & \text{if } m = 1; \\ (m, n, pg_2) & \text{if } m = 2, n \equiv 0 \pmod{24}, n \geq 0; \\ (m, n, pg_3) & \text{if } m = 3, n \equiv 8 \pmod{24}, n \geq 0; \\ (m, n, pg_4) & \text{if } m = 4, n \equiv 16 \pmod{24}, n \geq 0; \\ (m, n, p) & \text{otherwise.} \end{cases}$$

Let $c = (baba^{-1})^4$; then $(0, n, p)c = (0, n + 8(-1)^n, p)$, while $(m, n, p)c = (m, n, p)$ if $m \neq 0$. Let $b_i = a^i ba^{-i}$ and $\hat{g}_i = c^{i-2} b_i c^{-3} b_i c^{5-i}$ for $i = 2, 3, 4$; then $(m, n, p)\hat{g}_i = (0, 0, pg_i)$ if $m = n = 0$, while $(m, n, p)\hat{g}_i = (m, n, p)$ otherwise.

ACKNOWLEDGMENTS. I thank the editor and referee for their encouragement and advice.

REFERENCES

1. Fred Galvin, Generating countable sets of permutations, *Abstracts Amer. Math. Soc.* 13 (1992), 509.
2. Graham Higman, B. H. Neumann, and Hanna Neumann, Embedding theorems for groups, *J. London Math. Soc.* 24 (1949), 247–254.
3. F. Levin, Factor groups of the modular group, *J. London Math. Soc.* 43 (1968), 195–203.
4. B. H. Neumann, An essay on free products of groups with amalgamations, *Philos. Trans. Roy. Soc. London Ser. A* 246 (1954), 503–554.
5. B. H. Neumann and Hanna Neumann, Embedding theorems for groups, *J. London Math. Soc.* 34 (1959), 465–479.
6. W. R. Scott, *Group Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.
7. W. Sierpiński, Sur les suites infinies de fonctions définies dans les ensembles quelconques, *Fund. Math.* 24 (1935), 209–212.
8. Stan Wagon, personal communication, 1979.
9. Composition of functions (solution of Problem 6244, proposed by John Myhill), *Amer. Math. Monthly* 87 (1980), 676–678.

Department of Mathematics
University of Kansas
Lawrence, KS 66045-2142

Abelian Forcing Sets

Joseph A. Gallian and Michael Reid

Many readers of the MONTHLY have encountered particular cases of the following question. Suppose G is a group and n is an integer with the property that $(ab)^n = a^n b^n$ for all a and b in G . Which values of n imply that G is Abelian? Indeed, standard exercises in undergraduate abstract algebra textbooks ([1], [2], [3], [4]) are to show that $n = 2$ and $n = -1$ are two such values. Are there others? If $n \in \mathbb{Z}$, we say that a group G is n -Abelian if $(xy)^n = x^n y^n$ for all $x, y \in G$. Thus

our question may be reformulated as “for which integers n is an n -Abelian group necessarily abelian?” If p is any prime, consider the non-Abelian group

$$G_p = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{Z}_p \right\}.$$

If p is odd, then $x^p = e$ for all $x \in G_p$. We say that a group G has *exponent* n if $x^n = e$ for all $x \in G$. Thus, G_p has exponent p . Also, G_2 (which is isomorphic to the group of symmetries of a square) has exponent 4. Note that if G is a group with exponent n , then for any integer k , G is kn -Abelian and $(kn + 1)$ -Abelian. The examples G_p are now sufficient to show that the only integers n for which n -Abelian implies Abelian are $n = 2$ and $n = -1$. Indeed, for p odd, G_p is pk -Abelian and $(pk + 1)$ -Abelian for any integer k , while G_2 is $4k$ -Abelian and $(4k + 1)$ -Abelian.

More generally, let us call a set of integers T *Abelian forcing* if whenever G is a group with the property that G is n -Abelian for all n in T , then G is Abelian. So far we have seen that the only singleton Abelian forcing sets are $\{-1\}$ and $\{2\}$. What about other sets? Both of Herstein’s algebra textbooks ([3, p. 31] and [4, p. 57]) include the exercise that sets containing three consecutive integers are Abelian forcing. Moreover, one of Herstein’s books ([4, p. 57]) has an exercise that $\{3, 5\}$ is an Abelian forcing set. In contrast, the set $\{3, 7\}$ is not Abelian forcing, as G_3 is both 3-Abelian and 7-Abelian.

What characterizes the Abelian forcing sets? Although we could not find the answer to this precise question in the literature, some of the essential features of our argument below can be gleaned from a paper by F. Levi [5] written in the group-theoretic language of fifty years ago. (Levi investigated the question of when the mapping $a \mapsto a^n$ is a group endomorphism.) Our formulation of the question, the answer and the proof make the material more accessible to undergraduates.

Theorem. *A set T of integers is Abelian forcing if and only if the greatest common divisor of the integers $n(n - 1)$ as n ranges over T is 2. (Note that each $n(n - 1)$ is even.)*

Proof: The necessity of the condition again follows from the examples G_p . For p prime, let $T_p = \{n \in \mathbb{Z} \mid 2p \text{ divides } n(n - 1)\}$. Then for p odd, $T_p = \{pk, pk + 1 \mid k \in \mathbb{Z}\}$, while $T_2 = \{4k, 4k + 1 \mid k \in \mathbb{Z}\}$. From our earlier observation, G_p is n -Abelian for each $n \in T_p$, so T_p is not Abelian forcing. This proves necessity.

To prove sufficiency of the condition, suppose that $T \subseteq \mathbb{Z}$ satisfies $\gcd(n(n - 1) \mid n \in T) = 2$, and G is a group which is n -Abelian for all $n \in T$. Let $S = \{n \in \mathbb{Z} \mid G \text{ is } n\text{-Abelian}\}$, so that $T \subseteq S$. First note that if $m, n \in S$, then $mn \in S$. Also, if $n \in S$, then for any $x, y \in G$, we have $(xy)^n = x^n y^n$, so that $(yx)^{n-1} = x^{n-1} y^{n-1}$, whence $(yx)^{1-n} = y^{1-n} x^{1-n}$. Thus, if $n \in S$, then $1 - n \in S$. Since $n = 1 - (1 - n)$, the converse holds as well.

Our main difficulty at this point is that S is not closed under addition. However, suppose that $n \in S$ has the property that $x^n \in Z(G)$ (the center of G) for all $x \in G$. Then, for arbitrary $m \in S$, $x, y \in G$, we have $(xy)^{m+n} = (xy)^m (xy)^n = x^m y^m x^n y^n = x^m x^n y^m y^n = x^{m+n} y^{m+n}$, so $m + n \in S$.

This motivates the definition $R = \{n \in S \mid x^n \in Z(G) \text{ for all } x \in G\}$. It is easy to see that $n \in R$ if and only if $-n \in R$. Thus, from our previous remark, R is an additive subgroup of \mathbb{Z} . We now claim that if $n \in S$, then $n(n - 1) \in R$. We do this in several steps.

Note that if $n \in S$, then $1 - n \in S$, so $n(1 - n) \in S$. For arbitrary $x, y \in G$, we have $yx^n y^n y^{-1} = y(xy)^n y^{-1} = (yx)^n = y^n x^n$, so that $y^{1-n} x^n = x^n y^{1-n}$. Thus n -th powers commute with $(1 - n)$ -th powers. Now, for any $x \in G$, $x^{n(1-n)}$ is both an n -th power and a $(1 - n)$ -th power. Thus, for any $y \in G$, $x^{n(1-n)}$ commutes with both y^n and y^{1-n} , and therefore also with y . This shows that $x^{n(1-n)} \in Z(G)$, so that $n(1 - n)$ and thus, also, $n(n - 1)$ are in R .

We are now in position to prove sufficiency. Since the greatest common divisor of the numbers $n(n - 1)$ for $n \in T$ is 2, the additive subgroup R of Z contains 2. Therefore, G is 2-Abelian, and thus Abelian. This proves sufficiency.

Finally, to see that $\{n, n + 1, n + 2\}$ is Abelian forcing, note that $n(n - 1) - 2(n + 1)n + (n + 2)(n + 1) = 2$, so that

$$\gcd(n(n - 1), (n + 1)n, (n + 2)(n + 1)) = 2.$$

The authors recently discovered that the problem addressed here appeared as a problem in the MONTHLY in 1974 (E2411, vol. 81, page 410). It is also a special case of a result of L. C. Kappe, "On n -Levi groups", *Arch. Math.*, 47 (1986) 198–210.

REFERENCES

1. J. B. Fraleigh, *A First Course in Abstract Algebra*, 4th ed., Addison-Wesley, Reading, MA, 1989.
2. J. A. Gallian, *Contemporary Abstract Algebra*, 2nd ed., D.C. Heath, Lexington, MA, 1990.
3. I. N. Herstein, *Topics in Algebra*, 2nd ed., John Wiley & Sons, New York, 1975.
4. I. N. Herstein, *Abstract Algebra*, 2nd ed., Macmillan, New York, 1990.
5. F. W. Levi, Notes on group theory. I, II. *J. Indian Math. Soc.* (N.S.), 8 (1944) 1–9.

Department of Mathematics
University of Minnesota, Duluth
Duluth, MN 55812
jgallian@ua.d.umn.edu

Department of Mathematics
University of California, Berkeley
Berkeley, CA 94720
reid@math.berkeley.edu

UNSOLVED PROBLEMS

Edited by: **Richard Guy**

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

Is There a k -Anisohedral Tile for $k \geq 5$?

John Berglund

Let T be a **monohedral** (all tiles are congruent) tiling of the Euclidean plane [2, p. 20]. Let $S(T)$ be the group of symmetries which map T onto itself. For a given tile T in T let the transitivity class of T be the collection of all tiles to which T can be mapped by one of the symmetries of $S(T)$. If T has precisely k transitivity classes, call T **k -isohedral**. Since a picture is worth 10^3 words, let us look at an example. FIGURE 1 is a tiling that is 1-isohedral. 1-isohedral tilings are also called merely isohedral. FIGURE 2 is a tiling that is 2-isohedral. The shaded tiles form one transitivity class, and the unshaded tiles form another.

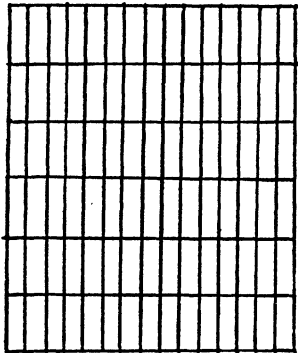


Figure 1. 1-isohedral tiling.

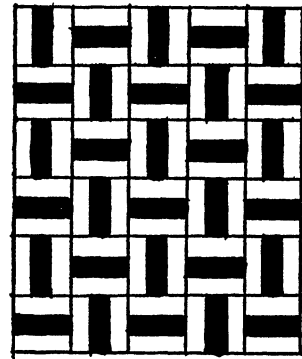


Figure 2. 2-isohedral tiling.

If a tile permits a k -isohedral tiling but not any n -isohedral tiling for $n < k$, call the tile **k -anisohedral**. For example, the tile given in FIGURE 3 is 2-anisohedral. The members of the shaded transitivity class bite the chins of the members of the unshaded transitivity class. The members of the unshaded transitivity class bite the noses. The tile in FIGURE 2 is not 2-anisohedral since it allows the 1-isohedral

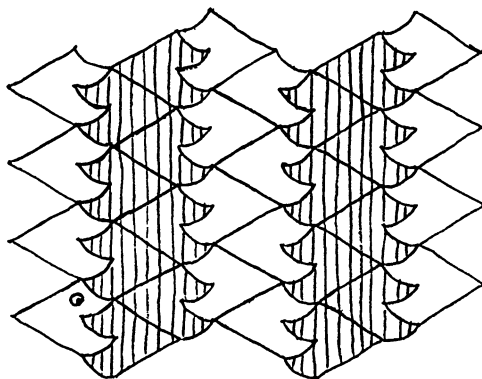


Figure 3. 2-anisohedral tile.

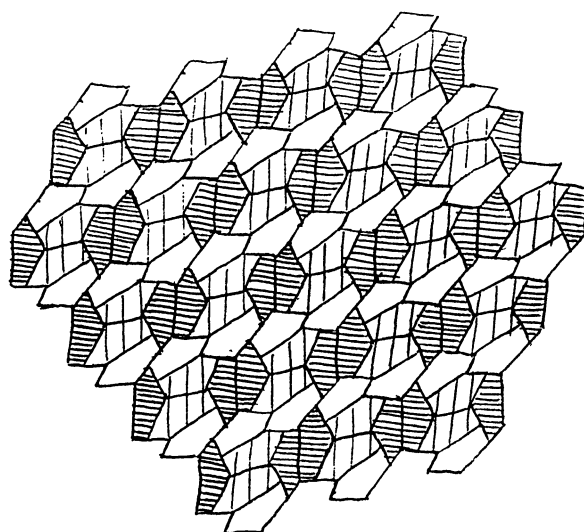


Figure 4. 3-anisohedral tile.

tiling given in FIGURE 1. 3-anisohedral tiles are rarer. As one example, take the Stein pentagon (FIGURE 4) [2, p. 518].

4-anisohedral tiles are still rarer. The figure shown in FIGURE 5 seems to be the first published example. Note that the shape is made by joining nine equilateral triangles at the edges.

Problem 1. Do there exist k -anisohedral tiles for every k ? Examples have been found for $k \leq 4$.

Problem 2. Characterize all k -anisohedral tiles for low values of k . This has been solved for $k = 1$ in Grünbaum and Shephard [1]. The general problem has not been solved even for $k = 2$.

To give the reader a taste of these topics, two 4-isohedral tilings are given in FIGURES 6 and 7. Are these shapes 4-anisohedral or not? The shape in FIGURE 6 is

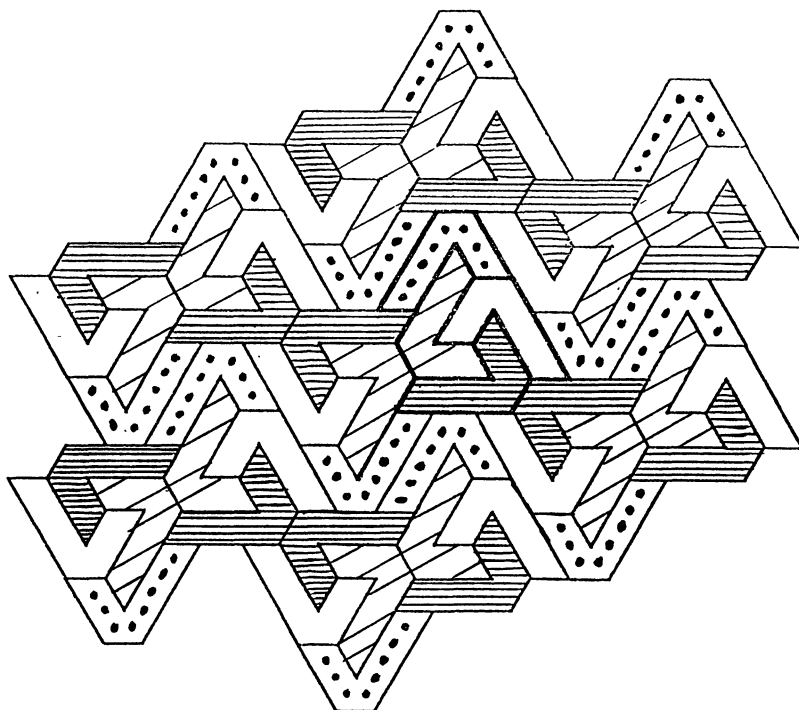


Figure 5. 4-anisohedral polyiamond.

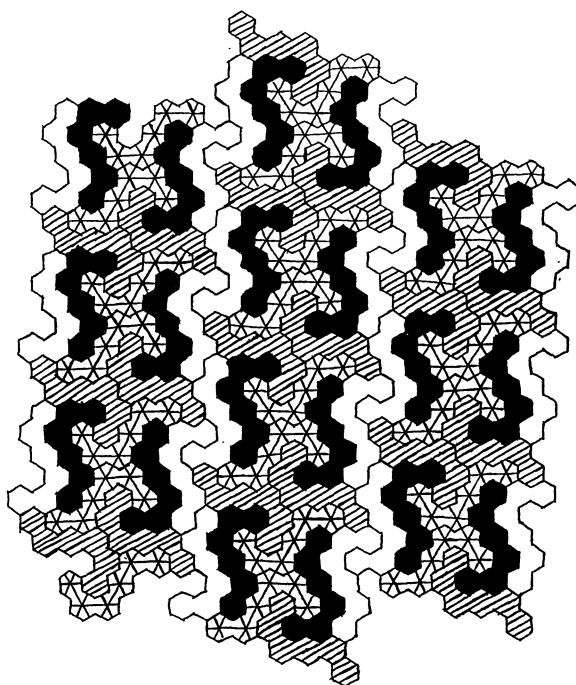


Figure 6. Hexahex in a 4-isohedral tiling.

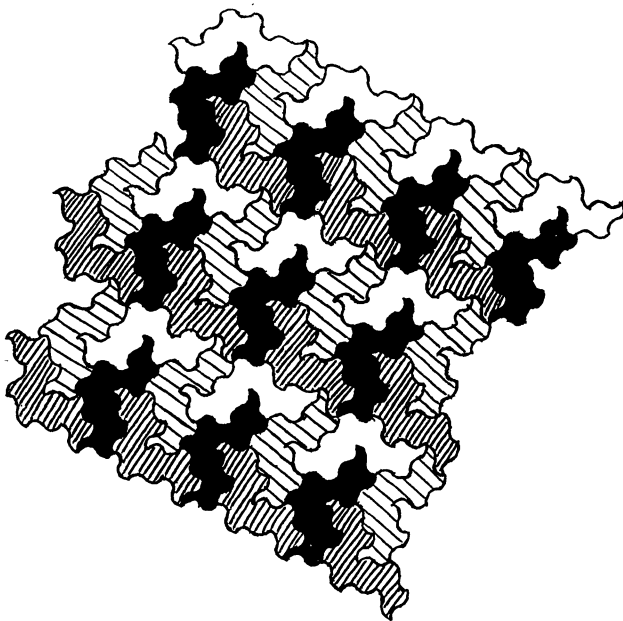


Figure 7. Modified 9-diamond in a 4-isohedral tiling.

made by joining six regular hexagons. The shape in FIGURE 7 is based on a shape made of nine equilateral triangles joined edge to edge; but the edges have been replaced with centrosymmetric curves.

REFERENCES

1. B. Grünbaum and G. C. Shephard, The 81 types of isohedral tilings of the plane. *Math. Proc. Cambridge Philos. Society* 82 (1977), 177–196.
2. B. Grünbaum and G. C. Shephard, *Tilings and Patterns*. Freeman and Co., New York, 1986.

Indiana Academy
Cicero, IN 46034

“...She knew only that if she did or said thus-and-so, men would unerringly respond with the complimentary thus-and-so. It was like a mathematical formula and no more difficult, for mathematics was the one subject that had come easy to Scarlett in her school-days.”

From *Gone With the Wind*
by Margaret Mitchell
Submitted by Steven C. Althoen

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before November 30, 1993 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10314. *Proposed by Andrew Vince, University of Florida, Gainesville, FL.*

Let b be an integer greater than 1. Let S be a set of integers containing 0 such that no two members of S are congruent modulo b . If

$$\sum_{i=1}^{\infty} \frac{s_i}{b^i} = 0,$$

with $s_i \in S$, prove that all $s_i = 0$.

10315. *Proposed by Mowaffaq Hajja, Yarmouk University, Irbid, Jordan.*

Let A and B be matrices with integer entries of sizes r by n and n by r , respectively, with $r < n$. Suppose that AB is an r by r identity matrix. Show that A can be enlarged to an n by n integral matrix having an integral inverse.

10316. *Proposed by Richard K. Guy, University of Calgary, Calgary, Alberta, Canada, and Richard J. Nowakowski, Dalhousie University, Halifax, N.S., Canada.*

For what pairs of integers a, b does ab exactly divide $a^2 + b^2 + 1$?

10317. Proposed by Juan Bosco Romero Márquez, Universidad de Valladolid, Valladolid, Spain.

Let $\triangle ABC$ be inscribed in a circle \mathcal{C} and let A', B', C' be the midpoints of the arcs $\widehat{BC}, \widehat{CA}, \widehat{AB}$, respectively.

(a) Prove that the incenter of $\triangle ABC$ is the orthocenter of $\triangle A'B'C'$.

(b) Prove that the pedal triangle of $\triangle A'B'C'$ is homothetic to $\triangle ABC$.

10318. Proposed by William P. Wardlaw, United States Naval Academy, Annapolis, MD.

Suppose that A is an n by n matrix with rational entries whose multiplicative order is 15; i.e. $A^{15} = I$, an identity matrix, but $A^k \neq I$ for $0 < k < 15$. For which n can one conclude from this that

$$I + A + A^2 + \cdots + A^{14} = 0?$$

10319. Proposed by Nick MacKinnon, Winchester College, Winchester, U.K.

Define

$$S_k(n) = \sum_{r=1}^n \sin(r^k).$$

Examination of graphs of $S_k(n)$ as a function of n for $1 < k < 2$ reveals some striking patterns. For example, when $k = 1.4$ the graph divides into clearly defined regions: between $n = 1$ and $n = 36$, one has $-.5 < S_{1.4}(n) < 3$; then the value of the function changes rapidly from $S_{1.4}(36) \approx 2.95$ to $S_{1.4}(49) \approx -5.7$; then one has $-6.2 < S_{1.4}(n) < -1.4$ as n goes from 49 to 225; then there is a rapid increase from $S_{1.4}(225) \approx -6.2$ to $S_{1.4}(257) \approx 15.7$. This pattern persists as far as graphs have been drawn.

(a) As a first step to understanding this phenomenon, determine the locations of the jumps in $S_{1.4}(n)$.

(b) Show that similar behavior may be expected for all k with $1 < k < 2$, with the flat regions being longer for k close to 1 and shorter for k close to 2.

10320. Proposed by Ignacy I. Kotlarski, Oklahoma State University, Stillwater, OK.

Under the assumption that f_0 , f_1 and f_2 are defined on $[0, \infty)$, Laplace transformable, and not equivalent to zero, solve the integral equation

$$\int_0^{\min(x_1, x_2)} f_0(x) f_1(x_1 - x) f_2(x_2 - x) dx = e^{-\max(x_1, x_2)} (1 - e^{-\min(x_1, x_2)}),$$

with $x_1 \geq 0$ and $x_2 \geq 0$, for the three functions f_0 , f_1 and f_2 .

10321. Proposed by Carl Axness, Sandia National Laboratories, Albuquerque, NM, Reinhard Schäfke, University of Essen, Essen, Germany, and David Arterburn, New Mexico Tech., Socorro, NM.

Let μ be a positive real number. Prove

$$\lim_{x \rightarrow 1^+} (\ln x)^{1/\mu} \sum_{i=1}^{\infty} x^{-(2i-1)\mu} = \frac{\Gamma(1/\mu)}{2\mu}.$$

NOTES

Notes: (10317) The *incenter* of a triangle is the center of its inscribed circle, and the *orthocenter* is the point of intersection of its altitudes. The feet of the altitudes of a triangle are the vertices of its *pedal triangle*. **(10319)** See the cover!

SOLUTIONS

Repeated Cyclotomic Factors

E 3442 [1991, 438]. *Proposed by Ray Wylie, Furman University, Greenville, SC.*

Given a sequence $\{b_n\}_{n=0}^{\infty}$ of real numbers such that $b_n = 0$ for n sufficiently large, put $B_s = \sum_{t=0}^{\infty} b_{mt+s}$ ($s = 0, 1, \dots, m-1$), and let us say that the sequence $\{b_n\}_{n=0}^{\infty}$ has property P_m , where m is a positive integer, if

$$B_0 = B_1 = \dots = B_{m-1}.$$

Suppose a_1, a_2, \dots, a_r are positive integers, not necessarily distinct, and let $C(n)$ be the number of r -tuples (n_1, n_2, \dots, n_r) of integers such that

$$n = n_1 + n_2 + \dots + n_r, \quad 0 \leq n_i \leq a_i \quad \text{for } i = 1, 2, \dots, r.$$

Prove that the k sequences

$$\{C(n)\}_{n=0}^{\infty}, \{nC(n)\}_{n=0}^{\infty}, \dots, \{n^{k-1}C(n)\}_{n=0}^{\infty}$$

all have property P_m if and only if at least k of the integers a_1, a_2, \dots, a_r are congruent to -1 modulo m .

Solution by O. P. Lossers, Eindhoven University of Technology, Eindhoven, The Netherlands. We consider $m > 1$, and let $\zeta = \exp(2\pi i/m)$. Let $B(x) = \sum_{n=0}^{\infty} b_n x^n$ and $F(x) = \sum_{j=0}^{m-1} B_j x^j$. Observe that $B(\zeta^t) = F(\zeta^t)$ for each $t \in \{1, 2, \dots, m-1\}$. Therefore, if $\{b_n\}_{n=0}^{\infty}$ has property P_m , then $B(\zeta^t) = 0$ for each $t \in \{1, 2, \dots, m-1\}$ and $B(x)$ is divisible by $1 + x + \dots + x^{m-1}$. Furthermore, if $B(\zeta^t) = 0$ for each $t \in \{1, 2, \dots, m-1\}$, then since $F(x)$ is a polynomial of degree $m-1$, $F(x)$ must be a constant times $1 + x + \dots + x^{m-1}$ so that $\{b_n\}_{n=0}^{\infty}$ has property P_m .

Observe that the polynomial $C(x) = \sum_{n=0}^{\infty} C(n)x^n$ satisfies

$$C(x) = \prod_{i=1}^r (1 + x + \dots + x^{a_i}).$$

Also, for every integer $t \geq 0$, the polynomial $\sum_{n=0}^{\infty} n^t C(n)x^n$ is a linear combination of the polynomials

$$C(x), xC'(x), x^2C''(x), \dots, x^t C^{(t)}(x).$$

Hence, if the k given sequences all have property P_m , then the previous paragraph implies that all the polynomials $C(x), C'(x), \dots, C^{(k-1)}(x)$ have at least one factor $1 + x + \dots + x^{m-1}$ which implies that $C(x)$ is divisible by $(1 + x + \dots + x^{m-1})^k$. Hence, ζ is a zero of $C(x)$ of multiplicity at least k . Since each factor $1 + x + \dots + x^{a_i}$ has only simple zeroes, it follows that $a_i + 1$ is a multiple of m for at least k values of i . For the converse, observe that if $a_i + 1$ is a multiple of m for at least k values of i , then $C(x)$ is divisible by $(1 + x + \dots + x^{m-1})^k$. We can conclude that ζ^t is a root of $\sum_{n=0}^{\infty} n^j C(n) x^n$ for each $j \in \{0, 1, \dots, k-1\}$ and each $t \in \{1, 2, \dots, m-1\}$. The converse now follows from the previous paragraph.

Solved also by D. Callan, R. J. Chapman (United Kingdom), M. Dindos (Slovakia), K. S. Kedlaya (student), N. Komanda, and the National Security Agency Problems Group. One incorrect solution was received.

Hardy's Inequality for Geometric Series

6663 [1991, 559]. *Proposed by Walther Janous, Ursulinengymnasium, Innsbruck, Austria and the editors.*

Show that

$$\sum_{j=1}^N \left(\frac{1 + x + x^2 + \dots + x^{j-1}}{j} \right)^2 < (4 \log 2)(1 + x^2 + x^4 + \dots + x^{2N-2})$$

for $0 < x < 1$ and all positive integers N ; also show that the constant $4 \log 2$ is best possible. (If we drop the factor $\log 2$, we have a special case of Hardy's inequality; see Hardy, Littlewood, and Pólya, *Inequalities*, pp. 239–242.)

Solution by Rolf Richberg, RWTH Aachen, Aachen, Germany. Observing that for $0 < x < 1$ and $N \in \mathbb{N}$

$$\int_x^1 \int_x^1 \frac{1 - (st)^N}{1 - st} ds dt = \sum_{j=1}^N \int_x^1 \int_x^1 (st)^{j-1} ds dt = \sum_{j=1}^N \left(\frac{1 - x^j}{j} \right)^2$$

we may reformulate the assertion as

$$\int_x^1 \int_x^1 \frac{1 - (st)^N}{1 - x^{2N}} \frac{ds dt}{1 - st} < (4 \log 2) \frac{1 - x}{1 + x} \quad (0 < x < 1, N \in \mathbb{N}), \quad (1)$$

which obviously would follow from

$$\int_x^1 \int_x^1 \frac{ds dt}{1 - st} < (4 \log 2) \frac{1 - x}{1 + x} \quad (0 < x < 1). \quad (2)$$

In order to prove (2) we consider the double integral:

$$\begin{aligned} \int_x^1 \int_x^1 \frac{ds dt}{1 - st} &= \int_x^1 \left[-\frac{1}{t} \log(1 - st) \right]_{s=x}^{s=1} dt \\ &= -\int_x^1 \log(1 - t) \frac{dt}{t} + \int_{x^2}^x \log(1 - t) \frac{dt}{t} \\ &= -2 \int_x^1 \log(1 - t) \frac{dt}{t} + \int_{x^2}^1 \log(1 - t) \frac{dt}{t}. \end{aligned}$$

Substituting $t = \tau^2$ in the last integral and noting $\log(1 - \tau^2) = \log(1 - \tau) + \log(1 + \tau)$, we obtain

$$\int_x^1 \int_x^2 \frac{ds dt}{1 - st} = 2 \int_x^1 \log(1 + t) \frac{dt}{t},$$

thus transforming (2) into

$$\int_x^1 \log(1 + t) \frac{dt}{t} < (2 \log 2) \frac{1 - x}{1 + x} \quad (0 < x < 1). \quad (3)$$

Now, differentiating with respect to t shows that $(1/t)\log(1 + t)$ is a decreasing and $(1 + (1/t))\log(1 + t)$ an increasing function of $t \in (0, 1)$. For $0 < x < 1$ we therefore have

$$\begin{aligned} \int_x^1 \log(1 + t) \frac{dt}{t} &< (1 - x) \frac{\log(1 + x)}{x} \\ &= \frac{1 - x}{1 + x} \left(1 + \frac{1}{x}\right) \log(1 + x) < \frac{1 - x}{1 + x} 2 \log 2, \end{aligned}$$

which proves (3).

Suppose (1) is valid with $4 \log 2$ replaced with a constant c . Taking the limit for $N \rightarrow \infty$ we then get

$$2 \int_x^1 \log(1 + t) \frac{dt}{t} = \int_x^1 \int_x^1 \frac{ds dt}{1 - st} \leq c \frac{1 - x}{1 + x} \quad (0 < x < 1).$$

In

$$\frac{2}{1 - x} \int_x^1 \log(1 + t) \frac{dt}{t} \leq \frac{c}{1 + x}.$$

(which comes from the transformation used to obtain (3)) let x tend to 1. It follows that $2 \log 2 \leq c/2$, i.e., $c \geq 4 \log 2$. Thus, the constant $4 \log 2$ is best possible.

Solved also by H. Morris.

An Extremal Set Problem

E3459 [1991, 754]. *Proposed by Constantin Adrian, Timisoara, Romania.*

Suppose X is an n -element set, $n \geq 12$, and suppose F is a family of 4-element subsets of X such that the intersection of each pair of distinct sets in F has at most two elements. Prove that there is a subset S of X containing at least $(6n - 6)^{1/3}$ elements such that none of the 4-element subsets of S is in the family F .

Solution by Fred Galvin, University of Kansas, Lawrence, KS. Call a subset of X *independent* if it contains no member of F . Assuming $n \geq 3$, we show that every maximal independent subset of X has size greater than $(6n)^{1/3}$.

Let S be a maximal independent set, with $k = |S|$. Clearly $k \geq 3$. Since S is maximal, for each $x \in X - S$ there is a 3-element set $f(x) \subseteq S$ such that $f(x) \cup \{x\} \in F$. The condition on F implies that f is injective. Hence $n - k \leq \binom{k}{3}$, or $6n \leq 6\binom{k}{3} + 6k = k^3 - 3k^2 + 8k \leq k^3 - 3$, and so $k \geq (6n + 3)^{1/3}$.

Editorial comment. Solvers proved various inequalities of the form $k \geq (6n + c)^{1/3}$ for $n \geq n_0$; by choosing n_0 large enough, c can be made arbitrarily large.

Fred Galvin noted that this problem is a special case of a class of problems discussed (but not explicitly solved) by Paul Erdős in “Problems and results on graphs and hypergraphs: similarities and differences,” in *Mathematics of Ramsey Theory* (J. Nešetřil and V. Rödl, eds.), Springer-Verlag, 1990, 12–28. The function $h_r(n, p, q)$ is the maximum integer m such that if each r -element subset of an n -set X is colored red or blue, then there exists a p -element subset of X containing at least q red r -sets or an m -element subset of X whose r -sets are all blue. The statement of this problem is $h_4(n, 5, 2) \geq (6n - 6)^{1/3}$ for $n \geq 12$.

Solved also by G. Calinescu (student, Romania), R. J. Chapman (U. K.), J. R. Griggs, R. Jeurissen (the Netherlands), I. Kastanas, O. P. Lossers (The Netherlands), B. Peterson, P. Tracy, and the proposer.

Source-Even Orientations of Graphs

E 3462 [1991, 755]. *Proposed by J. J. Rotman, University of Illinois at Urbana, Champaign, IL.*

Prove that any connected simple graph with an even number of edges has an orientation (assignment of direction to each edge) such that the number of edges leaving each vertex is even.

Solution I by Richard Holzstager, The American University, Washington, DC. Suppose the edges of the graph are oriented to minimize the number of vertices that are sources of an odd number of edges. Since the total number of edges is even, there must be an even number of such vertices. If this even number is positive, find a path connecting two of these vertices (guaranteed by the graph being connected) and reverse the orientation of each edge in the path. This does not change the parity of edges leaving any intermediate vertex along the path, but it changes the endpoints from odd to even, contradicting minimality.

Solution II by Jerrold Grossman, Oakland University, Rochester, MI. It is well-known that every connected graph with an even number of edges can be decomposed into edge-disjoint copies of P_3 , the path containing two edges. (See, for example, exercise 8.21a in G. Chartrand and L. Lesniak, *Graphs and Digraphs* (Second edition), Wadsworth, 1986.) Given such a decomposition, we orient the edges of each P_3 from its center toward its endpoints. This orientation has an even number of edges leaving every vertex.

Editorial comment. (1) Many solvers noted that simplicity of the graph is not necessary; neither of the above proofs requires this assumption. (2) Many solvers also mentioned the following easy extension to connected graphs with an odd number of edges: For any vertex v in such a graph, there is an orientation such that the number of edges leaving v is odd and the number of edges leaving every other vertex is even. (3) F. Galvin, J. Conklin, and E. Stone proved that the number of orientations having the desired property is exactly 2^{m-n+1} , where m is the number of edges and n is the number of vertices. (4) F. Galvin offered an extension to infinite graphs: Let G be an infinite graph and let V_F be the set of vertices of finite degree. Then, for any mapping $p: V_F \rightarrow \{0, 1\}$, there is an orientation of G such that, for every vertex $v \in V_F$, the number of edges leaving v has the same parity as $p(v)$.

Solved also by 46 others and the proposer.

A Permutation on the Cube

6670 [1991, 862]. *Proposed by R. H. Jeurissen, Toernooiveld, Nijmegen, The Netherlands.*

Let $\{0, 1\}^n$ denote the set of n -bit strings of zeros and ones. If $(a_1, \dots, a_n) \in \{0, 1\}^n$, let $\pi_n(a_1, \dots, a_n)$ be the string (b_1, \dots, b_n) given by $b_1 = a_1$ and $b_k \equiv a_k + a_{k-1} \pmod{2}$ for $1 < k \leq n$. Since (a_1, \dots, a_n) can be retrieved from (b_1, \dots, b_n) , it is clear that π_n is a permutation of $\{0, 1\}^n$. Determine the cycle structure of the permutation π_n , i.e., the lengths of the cycles that occur and the number of cycles of each length.

Solution by Thomas Honold, Technische Universität München, München, Germany and Sonja Maus, Bonn, Germany (independently). The cycle lengths are powers of 2. If c_k denotes the number of cycles of length 2^k , then

$$c_k = \begin{cases} 2 & \text{if } k = 0 \\ 2^{-k}(2^{2^k} - 2^{2^{k-1}}) & \text{if } 2 \leq 2^k \leq n \\ 2^{-k}(2^n - 2^{2^{k-1}}) & \text{if } n \leq 2^k \leq 2n \\ 0 & \text{if } 2^k \geq 2n \end{cases}$$

To establish this result, identify $\{0, 1\}^n$ with the vector space V of n -tuples over $GF(2)$, the field with two elements. Then π_n is an automorphism of V , and $\pi_n = I + N$, where I is the identity operator and N is the shift operator defined by $N(a_1, \dots, a_n) = (0, a_1, \dots, a_{n-1})$. Note that N is nilpotent, with $N^n = 0$. Since we are working over $GF(2)$, we have $\pi_n^{2^k} = I + N^{2^k}$ for every integer positive k . It follows that the order of π_n is a power of 2 and hence so is the length of every cycle of π_n . Now $v \in V$ is fixed by π_n if and only if $v \in \ker(N)$, and v lies in a cycle of length 2^k with $k \geq 1$ if and only if $v \in \ker(N^{2^k}) - \ker(N^{2^{k-1}})$. The result now follows from the observation that $\dim(\ker(N^j)) = \min\{n, j\}$.

Editorial comment. J. C. Binz and Tad White (independently) extended the result to the analogous automorphism of the vector space of n -tuples over $GF(p)$ for any prime p . Arthur Woerheide extended it even further to the analogous automorphism of G^n where G is any finite-dimensional vector space over $GF(p)$. These extensions are easily derived by appropriate modifications to the given solution.

Solved by 30 solvers (including those cited) and the proposer.

Deferred Cesàro Means

10217 [1992, 362]. *Proposed by Brian Philp, The University of Birmingham, Birmingham, England.*

Suppose $\{\alpha_j\}_{j=1}^\infty$ is a sequence of complex numbers.

- (a) Prove that if $n^{-1} \sum_{j=n}^{2n} \alpha_j \rightarrow \lambda$ and $n^{-1} \sum_{j=n}^{4n} \alpha_j \rightarrow 3\lambda$, then $n^{-1} \alpha_n \rightarrow 0$.
 (b) Is it true that if $n^{-1} \sum_{j=n}^{3n} \alpha_j \rightarrow 2\lambda$ and $n^{-1} \sum_{j=n}^{9n} \alpha_j \rightarrow 8\lambda$, then $n^{-1} \alpha_n \rightarrow 0$?

Solution by Robin J. Chapman, University of Exeter, Exeter, U. K.

(a) By replacing α_n by $\alpha_n + \lambda$ we may, and shall, assume that $\lambda = 0$. Let $\beta_n = n^{-1} \alpha_n$. Now

$$\frac{\alpha_{2n}}{n} = \frac{1}{n} \sum_{j=n}^{2n} \alpha_j + \frac{2}{2n} \sum_{j=2n}^{4n} \alpha_j - \frac{1}{n} \sum_{j=n}^{4n} \alpha_j$$

and so $\beta_{2n} \rightarrow 0$ as $n \rightarrow \infty$. It follows that $\delta_n = n^{-1} \sum_{j=n}^{2n-1} \alpha_n \rightarrow 0$ as $n \rightarrow \infty$. Now $(n+1)\delta_{n+1} - n\delta_n = \alpha_{2n+1} + \alpha_{2n} - \alpha_n$ and so

$$\begin{aligned}\beta_{2n+1} &= \frac{n+1}{2n+1} \delta_{n+1} - \frac{n}{2n+1} \delta_n - \frac{2n}{2n+1} \beta_{2n} + \frac{n}{2n+1} \beta_n \\ &= \gamma_n + \frac{n}{2n+1} \beta_n\end{aligned}$$

where $\gamma_n \rightarrow 0$ as $n \rightarrow \infty$. Hence $|\beta_{2n+1}| \leq |\gamma_n| + |\beta_n|/2$. Given $\varepsilon > 0$ choose N such that $|\gamma_n|, |\beta_{2n}| < \varepsilon$ if $n \geq N$. Let $K = \max(|\beta_N|, |\beta_{N+1}|, \dots, |\beta_{2N}|)$. I claim that if $2N \leq m \leq 4N$ then $|\beta_m| \leq \varepsilon + K/2$. This is trivial if m is even and if $m = 2n+1$ is odd then $|\beta_m| \leq |\gamma_n| + |\beta_n|/2 \leq \varepsilon + K/2$ as required. If we define $f_\varepsilon(K) = \varepsilon + K/2$ then iterating this argument gives $|\beta_m| \leq f_\varepsilon^r(K)$ (the r -fold iterate of f_ε applied to K) for $2^r N \leq m \leq 2^{r+1} N$. Now for fixed ε and large r , $f_\varepsilon^r(K) \rightarrow 2\varepsilon$ and hence $|\beta_m| \leq 3\varepsilon$, as required.

(b) The answer is “no.” I claim we can choose the α_{3k+1} for $k \geq 0$, arbitrarily so that $\sum_{j=n}^{3n} \alpha_j = \sum_{j=n}^{9n} \alpha_j = 0$. We make $\alpha_{3k} = 0$ for all $k \geq 1$ and if $k \geq 0$ define α_{3k+2} for $k \geq 0$, recursively by $\alpha_{3k+2} = -\sum_{j=k+1}^{3k+1} \alpha_j$. It is now clear that $\sum_{j=n}^{3n} \alpha_j = \sum_{j=n}^{9n} \alpha_j = 0$. But choosing say $\alpha_{3k+1} = k^2$ we can make $\{n^{-1}\alpha_n\}_{n=1}^\infty$ unbounded.

Editorial comment. The problem arose in connection with noticing that the claimed necessary and sufficient condition

$$\frac{1}{n} \sum_{j=n}^{\lambda n} \alpha_j \rightarrow (\lambda - 1)s \quad \text{as } n \rightarrow \infty \text{ for some } \lambda > 1$$

for Cesàro summability given as “analytic background” in Theorem IV of N. H. Bingham, “On Tauberian theorems in probability theory”, *Nieuw Arch. Wisk.* (4) 3 (1985), 157–166 was incorrect. Part (b) provides a counterexample to this statement. In general, a second value of $\lambda \geq 1$ is needed. A proof of the positive result given in part (a) due to Prof. B. Kuttner was provided by the proposer.

Solved also by M. Dindos (Slovakia), N. J. Fine, K. S. Kedlaya (student), O. P. Lossers (The Netherlands), M. Mócsy (Hungary), and R. Stong. One solution dealing only with part b, one for which only part b was correct, and one incorrect solution were also received.

Collaborating editors: David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Jerrold R. Griggs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttmann, Frank B. Miles, Richard Pfiefer, Stephen L. Portnoy, J. O., Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, Edward T. H. Wang, and William E. Watkins.

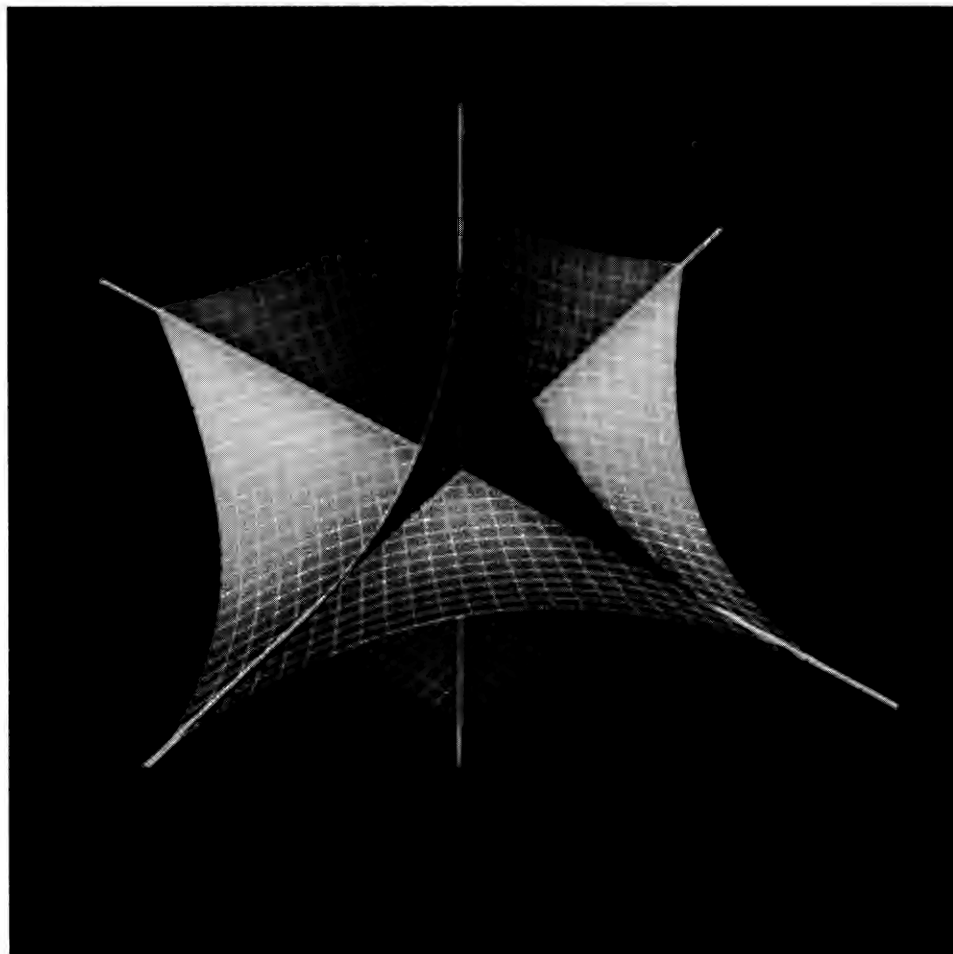
Answer to Picture Puzzle:
(p. 538)

John Littlewood, sometimes described as the name Hardy invented for a collaborator.

The American Mathematical Monthly



Volume 100, Number 7 / AUGUST-SEPTEMBER 1993



NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

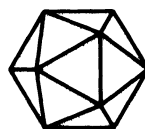
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International, Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 7 / AUGUST–SEPTEMBER 1993
(ISSN 0002-9890)



Contents

ARTICLES

Thomas Archer Hirst—Mathematician Xtravagant III. Göttingen and Berlin / J. HELEN GARDNER and ROBIN J. WILSON 619

Bisectors of Triangles and Tetrahedra / W. A. BEYER and BLAIR SWARTZ 626

More on Rectangles Tiled by Rectangles / D. G. MEAD and S. K. STEIN 641

Ramanujan—For Lowbrows / BRUCE C. BERNDT and S. BHARGAVA 644

Chebychev Polynomials and Regular Polygons / D. Y. SAVIO and E. R. SURYANARAYAN 657

Small-Group Learning / JULIAN WEISSGLASS 662

A Fast Pick-Type Approximation for Areas of H -Polygons / DING REN, KRZYSZTOF KOŁODZIEJCZYK, GRATTAN MURPHY, and JOHN REAY 669

FEATURES

COMMENTS 618

PICTURE PUZZLE 661

NOTES 674

COMPUTER SCIENCE SAMPLER

Zero-Knowledge Proofs / CATHERINE C. McGEOCH 682

THE AUTHORS 686

PROBLEMS AND SOLUTIONS 688

REVIEWS

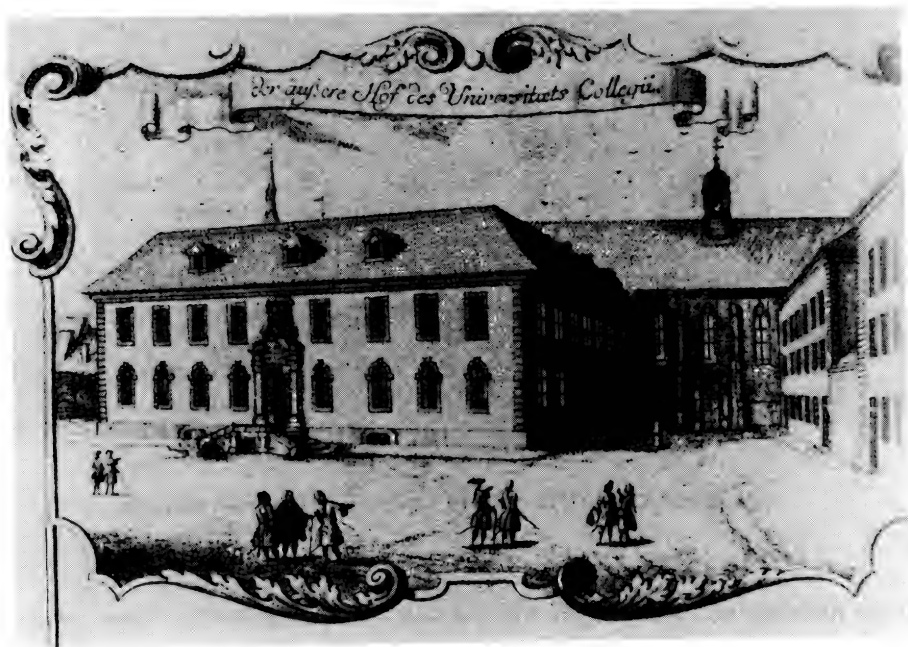
A First Course in Noncommutative Rings. By T.-Y. Lam / LANCE SMALL 698

TELEGRAPHIC REVIEWS 700

Thomas Archer Hirst— Mathematician Xtravagant III. Göttingen and Berlin

J. Helen Gardner and Robin J. Wilson

I carry with me all manner of letters of introduction to Göttingen, but the making of new acquaintances is ever a task to me; and for my own part I would rather have dispensed with extraneous help in doing so. It is a worse thing to be over- than under-estimated, and there is ever far more satisfaction in silently carving one's own path than in having it made ready and carpeted for us...



The University of Göttingen

With the completion of his Ph.D. thesis in Marburg, Thomas Hirst decided to travel, making Göttingen his first port of call. Here he spent two weeks at the University, attending lectures and conducting magnetic experiments with the physicist Wilhelm Weber, 'a curious little fellow [who] speaks in a shrill, unpleasant and hesitating voice'. He also attended Moritz Stern's 'beautifully clear' lectures on integral calculus and mechanics, and was much impressed by him.

6th August 1852: To-day I called on Weber again. He received me kindly and explained to me a new method of his of determining the inclination of the Magnet. He speaks and stutters on unceasingly; one has nothing to do but to listen. Sometimes he laughs for no earthly reason, and one feels sorry at not being able to join him. At 2 p.m. I went with him to make some experiments—he, I and another student made a determination of the Magnetic Inclination. I read off the “Auschlags Winkel” [angle of inclination] and though my first attempt, we got a result of $67^{\circ} 26'$ —that is within the daily variation . . .

Stern is a stern fellow, a firm, rub-against-able fellow—not an atom of unnecessary ceremony about him, but the greatest plainness and character. We got a bit of brown bread and butter together, as he would have taken himself, with a glass of water to it, and then a cigar. At first we talked about a dark point in his lecture, then on a multitude of topics. Stern is a widower with a family, a housekeeper; I believe his wife destroyed herself. At any rate, she went insane and either killed herself or died. It is said that for long after he was a misanthrope—one can see several deep wrinkles of sorrow in his face, though his manner is now gentle and quiet . . .



Carl Friedrich Gauss (1777–1855) on the terrace of Göttingen University

But the highlight of his Göttingen trip was a visit to Carl Friedrich Gauss, which he recorded in great detail.

12th August 1852: . . . Personally he is a venerable, fine old fellow, with a contented manly expression. There is an extraordinary aspect of power about him and his every word: without effort he suggests to every one the presence of manly might. He is about 80 years of age, but not a trace of superannuation is to be seen about him. He can even read without spectacles. Although our interview as far as the conversation was concerned was not brilliant or extraordinary, for in it there was no effort on either side, yet for remembrance sake I will try to relate it. No sooner was the first word spoken than I felt perfectly at ease, and he pointed for me to sit on the sofa and took a chair close by me. We spoke of course all in German, though he can speak English.

Tom. I am sorry, Professor, that I have not had the opportunity of hearing a lecture from you during my stay in Göttingen. It was, in fact, one of my principal motives for coming—a kind of

curiosity perhaps it may be called, yet for a lover of science I hope at least an excusable one. *Gauss*. Ah, this semester few students announced themselves, and at my age, with other work to be yet finished by me, and the hot summer before me, I was glad rather than otherwise to be dispensed from the task.

Tom (after a short silence). Have you ever been in England, Professor?

Gauss. No, I never got further than Belgium, and now the difficulties of the journey, as well as the change of life and habits render it impossible for me.

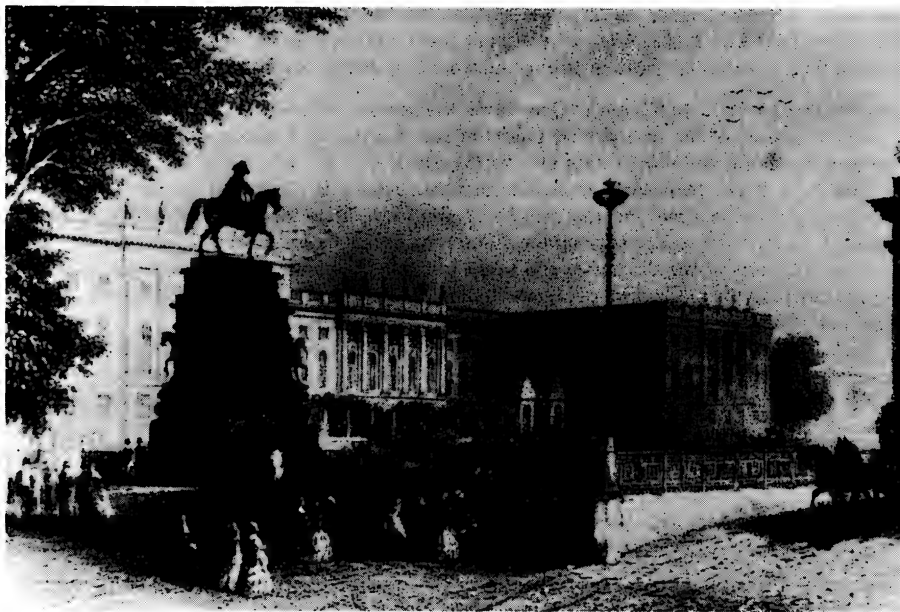
Tom. It is true this difference in habits and life affect even younger and stronger persons. I suffered somewhat myself from the same cause.

Gauss. Yet, considering the great esteem I have for the English as a nation, which we may consider a model for us as far as steady, persevering toil and firmness of character are concerned, it is now strange to me that in my younger days I never visited it.

Tom. You have, however, the consolation to know that it was your own work that prevented you. Yet there is something about your German life (especially student life) that so far excels English as yearly to draw more and more of us to your universities . . .

So we chatted on quite comfortably for three-quarters of an hour, and then I bid the old veteran good-bye and thanked him heartily. I left him copies of some of Tyndall's memoirs and of my own dissertation.

As a mathematician, Gauss is without doubt our Sir Isaac Newton. Perhaps no one ever had a firmer reliance in the absolute truth of mathematics, or lived more in it. It is to him the foundation of the universe on which God himself has built . . .



The University of Berlin, now known as Humboldt University

After leaving Göttingen, Hirst spent several weeks travelling around Germany and Austria with his brother John and some friends from Marburg. He then moved to Berlin, where he spent the winter semester, from October to April. On his arrival, he went to visit the algebraist Ferdinand Eisenstein, 'a young and highly promising mathematician'. Unfortunately, Eisenstein had died just the day before. This, understandably, upset Hirst considerably.

12th October 1852 ... A young, able fellow, cut down the moment he was making his ability known and useful—a fellow of deep intellect and great industry, too, as late journals can show. We entered joyfully, thinking to see him and know him; we left it awe-struck, silent and sad.

The next morning he called on Lejeune Dirichlet, the distinguished analyst and number theorist, and ‘met with a very hearty reception’.

13th October 1852: He is a rather tall, lanky-looking man, with moustache and beard about to turn grey (perhaps 45 years old), with a somewhat harsh voice and rather deaf: it was early, he was unwashed, and unshaved (what of him required shaving), with his “schlafrock”, slippers, cup of coffee and cigar... I thought, as we sat each at an end of the sofa, and the smoke of our cigars carried question and answer to and fro, and intermingled in graceful curves before it rose to the ceiling and mixed with the common atmospheric air, “If all be well, we will smoke our friendly cigar together many a time yet, good-natured Lejeune Dirichlet.”

Although he continued to read widely in all branches of mathematics, he increasingly felt the need to further his knowledge of geometry. He was particularly interested in the relationship between synthesis and analysis.

15th October 1852: After having purchased three valuable volumes, Carnot’s “Geometrie de Position”, Gauss’s “Theory of Numbers” (in French) and Cauchy’s application of Diff: Cal: to Geometry, I find my hands full of work. My own books have also arrived from Marburg; how dependent one is on books. I felt lost without them, and had to ask myself: “Tom, hast thou nothing then in thee, but must be strung and wound up before thou canst begin playing?”...

18th October 1852: ... In Carnot’s “Geometrie de Position” with which I am now engaged, after an able discussion of the comparative merits of and distinctions between so-called analysis and synthesis occurs the following rather noteworthy paragraph:

“Synthesis is not exclusively applied to mathematics—it is in general the art of reasoning with justice: whatever may be the subject of argument. It is identical with what is termed dialectics... Analysis proceeds generally differently, in that series of transformations are made on truncated parts of the discourse, and taken isolately are unintelligible, but which submitted like the others to the mechanism of argumentation can by a new series of transformations lead to clear and precise results—as much so, indeed, as those deduced synthetically...”

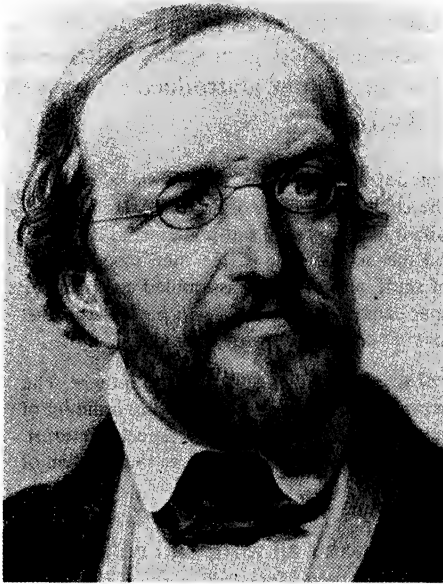
... Carlyle tells me that just at this time he [Carnot] was taking his part in the French Revolution—who knows but he may have written it after that memorable dinner party at which he and Robespierre were with others present, when Carnot slipped out of the room, searched Robespierre’s pocket that he had laid aside, and found therein a sentence of death for himself and others with whom that Robespierre had been chatting quite coolly...

It was not long before the lectures began. Hirst was particularly impressed by those of Jakob Steiner, both for his pure synthetic approach to geometry (which accorded with Hirst’s own views) and also for his attitudes to education, which he acquired while studying at the school of the Swiss reformer Pestalozzi.

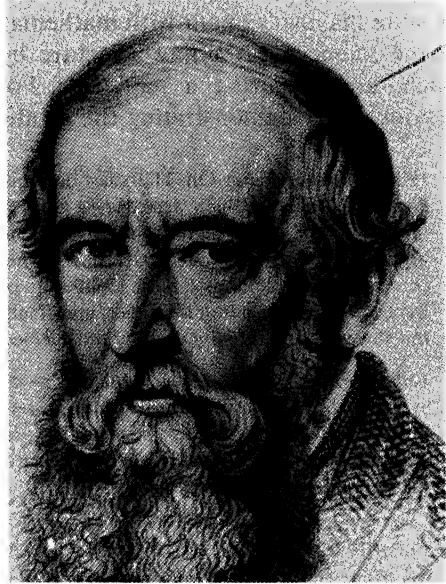
28th October 1852: ... I have heard Steiner twice, and am well pleased with him. He is a middle-aged man, of pretty stout proportions, has a long, intellectual face, with beard and moustache, and a fine prominent forehead, hair dark and rather inclining to turn grey. The first thing that strikes you on his face is a dash of care and anxiety almost pain, as if arising from physical suffering. Before starting he sets his chair right, looks all round, finds the window must be opened, and with difficulty gets started. Then in a short time he will ask them to close the window within a hand’s breadth, for he has rheumatism. All these point to physical nervous weakness. His Geometry is famed for its ingenuity and simplicity—he is an immediate pupil of Pestalozzi: in his youth was a poor shepherd boy, and now a professor. His argument is that the simplest way is the best; he tries ever to find out the way Nature herself adopts (not always, however, to be relied upon). Mathematics he defines to be the “Science of what is self-evident”...

But most of all, he admired the teaching of Dirichlet:

31st October 1852: Dirichlet cannot be surpassed for richness of material and clear insight into it: as a speaker he has no advantages—there is nothing like fluency about him, and yet a clear eye and understanding make it dispensable: without an effort you would not notice his hesitating speech. What is peculiar in him, he never sees his audience—when he does not use the black-board at which time his back is turned to us, he sits at the high desk facing us, puts his spectacles up on his forehead, leans his head on both hands, and keeps his eyes, when not covered with his hands, mostly shut. He uses no notes, inside his hands he sees an imaginary calculation, and reads it out to us—that we understand it as well as if we too saw it. I like that kind of lecturing.



Lejeune Dirichlet (1805–1859)



Jakob Steiner (1796–1863)

While enjoying the lectures of Steiner and Dirichlet, he was also developing their friendship. His social calls became increasingly frequent, and his diary entries around this time seem to alternate between the two. However, the two men could not have been more different, and this difference shines out of the words—the grumpy and ailing Steiner, with ‘a power of insight possessed by no other living geometer, perhaps’, and the genial Dirichlet, with whom he was becoming ‘on terms of perfect friendship’. But, for all this, he seems to have been fond of both of them, and any adverse comments were statements of fact, as he saw it, rather than condemnations.

7th November 1852: ...To-day I called on Prof. Riess of the “Königliche Academie der Wissenschaft” ... Riess is a delicate, good man, with clear, deep insight. I listened with great interest to his talk about Dirichlet, Jacobi and Steiner. He told me fully the relations on which the latter stands with them all, and truly it is unexplainable. Riess says his vulgarity has by them all been slightly borne in consideration of his undoubted genius. But that some time ago without provocation Steiner cut them all. The probable reason is that Steiner, naturally of a testy disposition, which has been increased, too, by bodily illness, feels himself slighted that he has been 33 years “Ausserordentliche” [Extraordinary] Professor. The reason is clear: firstly he does

not know Latin, and that among German professors is held as a necessity: 2nd he is so terribly one-sided on the question of Synthetical Geometry that as an examiner he would not be liked. The more I hear, the more I am determined to see him and study him for myself.

14th November 1852: ... Wednesday evening I spent with Dirichlet: saw Mrs Dirichlet again, found she was sister to Mendelssohn—she played me several of her brother's pieces, to which I listened with great willingness.

21st November 1852: On Tuesday I called again to see Steiner. He came to me first in his ante-room, when we had a little interesting talk on his system of Synthetical Geometry, and its relation to Analysis. The latter he would by no means annihilate, and pleads justly that heretofore it has but had too great pre-eminence to the detriment of Synthesis. I mentioned that I had in view to translate his work into English—the old fellow's indifference towards me has been somewhat relaxing before, and this was the finishing stroke ...

Despite his involvement with mathematics, Hirst continued to keep up his interests in the sciences, attending a lecture by the distinguished geologist Christian von Buch, and admiring a model of Foucault's pendulum, introduced two years previously for demonstrating the rotation of the Earth.

19th December 1852: On Thursday du Bois [Reymond] came to take me to the Königliche Academie. Old von Buch was reading a paper on the chalk formations of America. He is also a fine old fellow, with a stern iron face ... Steiner walked in with a very self-possessed, glum face, which relaxed into a smile and a bow as he passed me near the door. He then stood at one corner of the table, took a large, deliberate pinch of snuff, and eyed the assembled company ... Buch was sitting reading his paper, and finally Steiner came to a standstill with his back against the stove, took snuff very largely, and seemed to have no connection with anybody.

13th February 1853: Monday at Magnus's—a party of scientific ladies and gentlemen there. The evening was spent by Magnus shewing us several experiments. Magnus was shewing a model of Foucault's experiment with the pendulum; as he said, for Dirichlet's and my especial interest. The only remark the ladies made on the matter (and they always make some) was that motion of the pendulum was "sehr gracieuse."

His studies took most of his waking hours, but it wasn't all work and no play. Hirst took advantage of the crisp winter weather to pursue a 'favourite amusement'.

20th February 1853: This has been a winter's week indeed—10° below Zero in Baumer. Cabs has been almost entirely replaced by sledges, and the streets are in a continual tinkle, tinkle. I have for the first time in my life had a ride in one, on Wednesday, and very easy riding I found it. On Wednesday, and this afternoon, I had some skating for the first time in Germany—indeed, for many years. Swimming in summer and skating in winter are the two greatest physical enjoyments I indulge in, and I have a weakness for both. The scene on the ice here in the Thiergarten is novel and attractive to me. There are ladies by the score skating beautifully ... Prof. Mitscherlich's daughter was skating gracefully past me, and I, rogue that I am, was looking at her skates (and the ankles to which they were fastened) instead of my own, and I suffered for it. With a fearful crash I came on my rump. I did not break the ice, but I certainly left "my mark" there in the shape of an asterisk ...

But more often than not, Hirst was working hard, involved both with his mathematics, and also with his teachers—in particular, Dirichlet and Steiner. There are comments on their teaching styles, as well as frequent references to their personalities and activities.

20th February 1853: ... I will here record a few peculiarities in Steiner's lectures. He never prepares them beforehand, but follows ideas as they suggest themselves. He thus often stumbles or fails to prove what he wishes at the moment, and at every such failure he is sure to make some characteristic remark ... Dirichlet has also his peculiarities—one is of forgetting time; he pulls his watch out, finds it past three, and runs out *without even finishing the sentence*.

3rd April 1853: ...Steiner remains working at home until nearly 4 p.m.; he then dines at Schultz's Wein Keller—after dinner he takes a long walk until 6 or 7, returns home, works or sleeps until 11:30 or 12, and then comes to Schultz's again to drink a glass of grog and eat his supper. Here he shows his testiness most—the waiter he rated soundly for not bringing him his usual sort of wine—a person interrupted him whilst speaking, and got from him a stormy reprimand. After 1 a.m. we accompanied him home... Yesterday after dining again with Knoblauch, I was walking down the Linden and heard an unmistakeable voice (viz. Steiner's) calling my name. I turned with him, and we had a short walk. We spoke much on the old question of the relative claims of Analysis and Synthesis in Geometry, and I found him more liberal than ever before. As we returned, I asked if for once he would step into my lodgings and sit half an hour with me. He did so. And thus for once he paid a visit—a thing he seldom or never does...

24th April 1853: Wednesday evening we spent with Dirichlet... During the evening Prof. Hensel (husband to the late Fanny Hensel, another sister of Mendelssohn's) came. It was proposed that we should make the experiment of moving the table (Tisch rücken) which is now the subject of fashionable twaddle. Hensel had seen it the evening before and believed in it thoroughly. I and Dirichlet were thoroughly sceptical, Mrs D. was indifferent, and Dickinson and another gentleman were inclined to be believers. We sat for half an hour, each one placing both hands on the table and placing the little finger of his right hand on the little finger of his neighbour's left hand. The table was a pretty stout round one, with one leg rolling on three pulleys. It was easily movable by a single person. The experiment was totally unsuccessful... I believe the table would have moved this evening had we all been in a sufficient unanimity; one thought the table was leaning a little in one direction; directly two or three more are intent upon it so moving—look anxiously in that direction,—and unconsciously help it. The scientific men in Berlin almost all ridicule the idea. It deserves, however, a closer experiment.

Shortly afterwards, Hirst left Berlin for two months in Paris, where he attended the mathematics lectures of Joseph Liouville and Gabriel Lamé, before returning to England to take up a schoolteaching appointment at Queenwood College, in Hampshire. His time at Queenwood, his marriage, and his subsequent return to France, will be described in the next article.

ACKNOWLEDGMENTS. A typed version of the Thomas Hirst diaries is held at the Royal Institution in London, and quotations from the diaries appear here by courtesy of the Royal Institution. The diaries have been edited by W. H. Brock and R. M. MacLeod, and were published in microfiche by Mansell, London, in 1980.

The Open University
Milton Keynes MK7 6AA
England

Bisectors of Triangles and Tetrahedra

W. A. Beyer and Blair Swartz

1. INTRODUCTION. Our principal goal is to discuss and illustrate (in FIGURES 2, 3 and 4) the envelope of the planes that bisect a tetrahedron. To bisect, here, means to divide into two pieces of equal volume. The envelope of the hyperplanes that bisect a simplex in R^n and the envelope of the lines dividing a triangle into two pieces of fixed relative size are also considered (the last in FIGURE 5). These problems arose in work in numerical hydrodynamics dealing with approximating, within local second-order accuracy, a smooth boundary separating a black and white region in the plane, given discretely located gray values associated with a blurring of that interface (Swartz [17], exemplified at the end of §8 below). But the problems have a much older history in hydrostatics and naval architecture, as they are also connected with the orientation and stability of floating bodies. Favard's book [9] contains a satisfactory discussion of envelopes in R^2 and R^3 ; there are elementary discussions in Courant [7] or Courant and John [8].

2. A GENERAL DESCRIPTION. The envelope E of the planes that bisect a tetrahedron is homeomorphic to such traditional examples of closed, one-sided surfaces as the "Roman surface" of Steiner (see: Francis [10, pp. 83–86], Hilbert and Cohn-Vossen [12, pp. 303–4], and Spivak [15, pp. 20–1 and p. 34]) and the heptahedron (see Hilbert and Cohn-Vossen [12, pp. 302–3] and Jones [13]). Indeed, we shall see that the envelope E also consists of seven pieces—it is like a heptahedron whose faces have been pinched to tangency along its edges—but each piece is now part of the zero set of its own polynomial in three variables. In contrast, the Roman surface is the zero set of a single polynomial in three variables that is of total degree four.

3. THE ENVELOPE OF THE BISECTING LINES OF A TRIANGLE. Extending a homework problem in Thomas [18, p. 508, #61] or examples in Lamb [14, p. 232, Ex. 3], Greenhill [11, p. 190], or Bouasse [6, §253, p. 382], the following proposition summarizes the information about the envelope of the lines bisecting a triangle.

Proposition. *The envelope E of the set of all lines that bisect a given triangle T is a simple continuous closed curve lying completely inside T . It consists of three parts—the i th part being a segment of a hyperbola whose asymptotes include the two edges containing the i th vertex. Each segment joins continuously with its neighbor to form a cusp where the two are mutually tangent to an intervening median of T —and each of E 's three cusps has the same order of sharpness as the graph of $y = |x|^{1/2}$ near $(0, 0)$. The expression $b_0V_0 + b_1V_1 + b_2V_2$ for the i th part of the envelope ($i = 0, 1, 2$), in terms of barycentric coordinates (b_0, b_1, b_2) relative to the vertices V_0, V_1, V_2 of T , is independent of T and is given by the usual requirement $b_0 + b_1 + b_2 = 1$ together*

with the special requirements

$$8 \prod_{\substack{j=0 \\ j \neq i}}^2 b_j = 1 \quad \text{and} \quad 1/4 \leq b_j \leq 1/2 \quad \text{for } j \neq i; \quad (3.1)$$

for which $1/4 \leq b_i \leq 1 - 1/\sqrt{2}$. See FIGURE 1.

Proof: We first recall an instance of an envelope's hyperbolic segment. The rest then follows using properties of nonsingular affine transformations.

The area of the right triangle $T(x)$, associated with the two coordinate axes and the line tangent to the hyperbola $H := \{(t, 1/t), t > 0\}$, is 2, independent of the point $(x, 1/x)$ ($x > 0$) of tangency. This is a consequence of the following argument. $T(x)$ consists partly of the $x \times 1/x$ rectangle whose diagonal is the radius vector to the point of tangency. And, as the magnitude of the tangent line's slope is $1/x^2$, the interior of $T(x)$ outside this rectangle consists of two $x \times 1/x$ right triangles.

Consider also, now, the isosceles right triangle T_4 of area 4 with vertices $V_0 := (0, 0)$, $V_1 := (2\sqrt{2}, 0)$, and $V_2 := (0, 2\sqrt{2})$. The envelope of the hypotenuses of those $T(x)$ that are completely inside T_4 is a portion of the envelope of the bisecting lines for T_4 . It is also a segment S of H inside T_4 . The right-most point of S is the point $P_r := (\sqrt{2}, 1/\sqrt{2})$ since the tangent line to H here is also the median of T_4 that passes through $(0, \sqrt{2})$ and the vertex V_1 . By symmetry, S 's left-most point is $P_l := (1/\sqrt{2}, \sqrt{2})$; and the tangent to H there is also the median of T_4 through V_2 . Since the curvature of H exists on H but vanishes nowhere, S has the

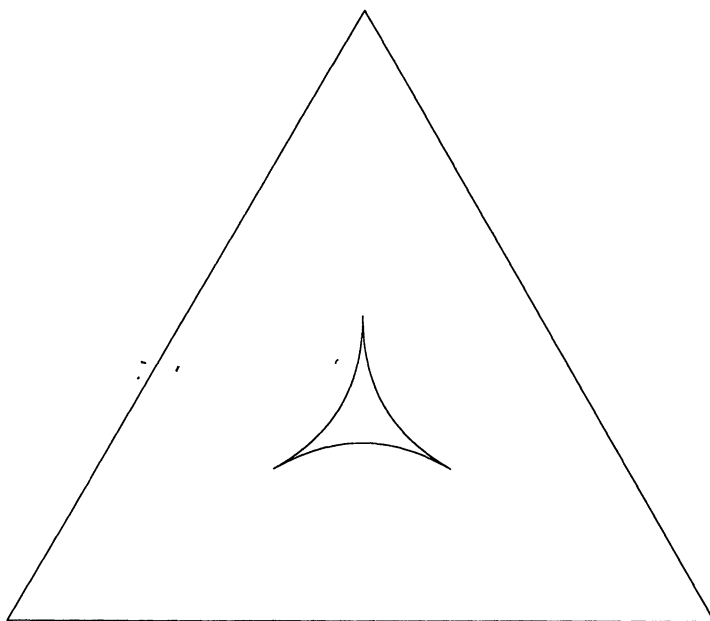


Figure 1. The bisecting envelope of a triangle: a line tangent to the cusped figure separates a corner from its opposite side and divides the triangle into two equal areas. The envelope's vertices are associated with the triangle's. Note that the normal line changes continuously along the whole of the envelope.

type of contact with these two tangent medians that is characteristic of a circle's contact with its tangent—this will be relevant to the cusp's order of sharpness.

Two of the barycentric coordinates of $(x, 1/x)$ with respect to the three vertices V_0, V_1 , and V_2 of T_4 are clearly $b_1 = x/(2\sqrt{2})$ and $b_2 = 1/(2x\sqrt{2})$. Consequently, on the segment S , b_1 and b_2 satisfy (3.1) with $i = 0$.

Now: a nonsingular affine map A maps tangent curves onto tangent curves and envelopes of curves onto envelopes of their images. Such a map A also maps lines onto lines, hyperbolas onto hyperbolas, non-degenerate triangles onto non-degenerate triangles and their medians onto medians, and (because A 's Jacobian determinant is independent of location) a line that bisects a triangle onto a line that bisects its image. Moreover, since a nonsingular affine map A is the sum of a linear map and a constant vector, the barycentric coordinates of a point V with respect to three points V_0, V_1 , and V_2 in general position (non-degenerate convex hull) are also the barycentric coordinates of $A(V)$ with respect to $A(V_0), A(V_1), A(V_2)$ (which are also in general position).

With this, proof of the remainder of the Proposition goes as follows. (a) Use three separate affine maps of T_4 onto an equilateral triangle T_3 (along with the symmetries of T_3) to show that the complete envelope for T_3 satisfies the Proposition (in particular, that neighboring hyperbolic segments are tangent at a common point on a median and so form the cusp as there characterized). (b) Then use a single affine map taking T_3 onto T .

4. PART OF THE ENVELOPE OF THE BISECTING HYPERPLANES OF A SIMPLEX IN n -DIMENSIONS. For $n > 2$ dimensions, consider the surface $H := \{x = (x_1, \dots, x_n) > 0 \text{ such that } f(x) = 0 \text{ with } f(x) := \prod_{j=1}^n x_j - 1\}$ (for three dimensions, see, e.g., Appell [1, p. 231, problem 11], Greenhill [11, p. 202], Struik [16, p. 73, problem 6], Courant and John [8, problem 8, p. 307]). Note that the j th component of the gradient ∇f of f on H is $(\nabla f)_j = 1/x_j$; hence points X in the hyperplane tangent to H at x satisfy $X \cdot \nabla f = n$. Thus the altitudes of the simplex $S(x)$ bounded by the coordinate hyperplanes and the hyperplane tangent to H at x are $(nx_j)_1^n$; so the volume of $S(x)$ is $n^n/n!$ independent of x . Restricting x so that $S(x) \subseteq S_2$:= the simplex of volume $2n^n/n!$ bounded by the coordinate hyperplanes and the hyperplane through and normal to $(1, \dots, 1)$ —that is, restricting the vertices of $S(x)$ to lie between those of S_2 and the origin—one can prove (see the Appendix) that an n -dimensional analog of the Proposition in §3 is:

Proposition. *Let $n \geq 2$. The envelope of the hyperplanes bisecting an n -dimensional simplex with vertices V_0, \dots, V_n consists partly of $n + 1$ hypersurfaces, each of degree n . The asymptotic hyperplanes of the i th of these hypersurfaces (some $0 \leq i \leq n$) are the hyperplane extensions of those n faces of the simplex that contain the i th vertex V_i . This i th hypersurface's $n + 1$ barycentric coordinates b_0, \dots, b_n with respect to V_0, \dots, V_n satisfy the usual requirement $\sum_{j=0}^n b_j = 1$, together with the relations*

$$2n^n \prod_{\substack{j=0 \\ j \neq i}}^n b_j = 1 \quad \text{and} \quad 1/(2n) \leq b_j \leq 1/n \quad \text{for } j \neq i; \quad (4.1)$$

for which $1/(2n) \leq b_i \leq 1 - (2^{-1/n})$. For $n = 3$ dimensions, see FIGURE 2.

The n -dimensional simplex has $n + 1$ vertices. The hyperplanes whose envelope is given by (4.1) separate the i th vertex from the n remaining vertices. Thus there are $n + 1$ such portions of the complete envelope of the simplex's bisecting

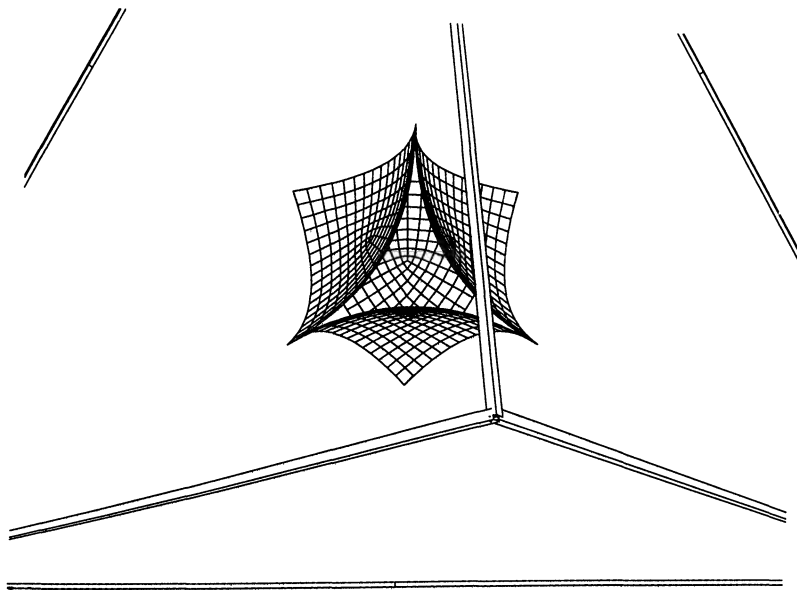


Figure 2. The corresponding bisecting envelope of a tetrahedron: a plane tangent to any of these four cupped, tri-edged surfaces separates a vertex from its opposite face and bisects the tetrahedron. It is now three *edges* of the envelope—the nearest coplanar hyperbolas—that are associated with the tetrahedron’s nearest vertex. There must be more bisecting planes.

hyperplanes. But there are other parts of the envelope. To obtain the remaining portions, the following outline of an algorithm can be carried out.

Let the $n + 1$ vertices be divided into two nonempty sets of vertices. Call the total number of such divisions $I(n)$ ($= (2^{n+1} - 2)/2$). For each such division, it is necessary to find the corresponding portion of the complete envelope. The complete envelope will then consist of $I(n)$ portions. In the next section we discuss the case $n = 3$; it is the only case for which we have a complete discussion.

5. THE ENVELOPE OF THE BISECTING PLANES OF A TETRAHEDRON.

A little thought concerning the equilateral tetrahedron ($n = 3$) in this context suggests that, corresponding to each non-intersecting pair of its edges, there should be a saddle-shaped surface bounded by four curves. One of these curves bounding a given saddle coincides with one of the three curves bounding a “cup” of FIGURE 2—each saddle is thereby connected to each of the four cups. The three saddles intersect (at the center of symmetry, for example) but are not tangent to each other—while they *are* tangent to the cups. FIGURES 3–4 illustrate these additional steps in the construction of the entire envelope of the bisecting planes of a regular tetrahedron. The captions explain the figures.

Further analysis is required to precisely specify these surfaces. As already noted, the affine invariance of the problem means that it suffices to determine the barycentric coordinates of the complete envelope E of the bisecting planes for any particular tetrahedron, and we shall fix on the tetrahedron U whose vertices are the origin $V_0 := \mathbf{0}$ and the three coordinate unit vectors $V_1 := \mathbf{i}$, $V_2 := \mathbf{j}$, and $V_3 := \mathbf{k}$. In other words, $U = \overline{\mathbf{0ijk}}$, where the overbar means “convex hull of.” In this context, if a portion of this envelope is a surface

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v);$$

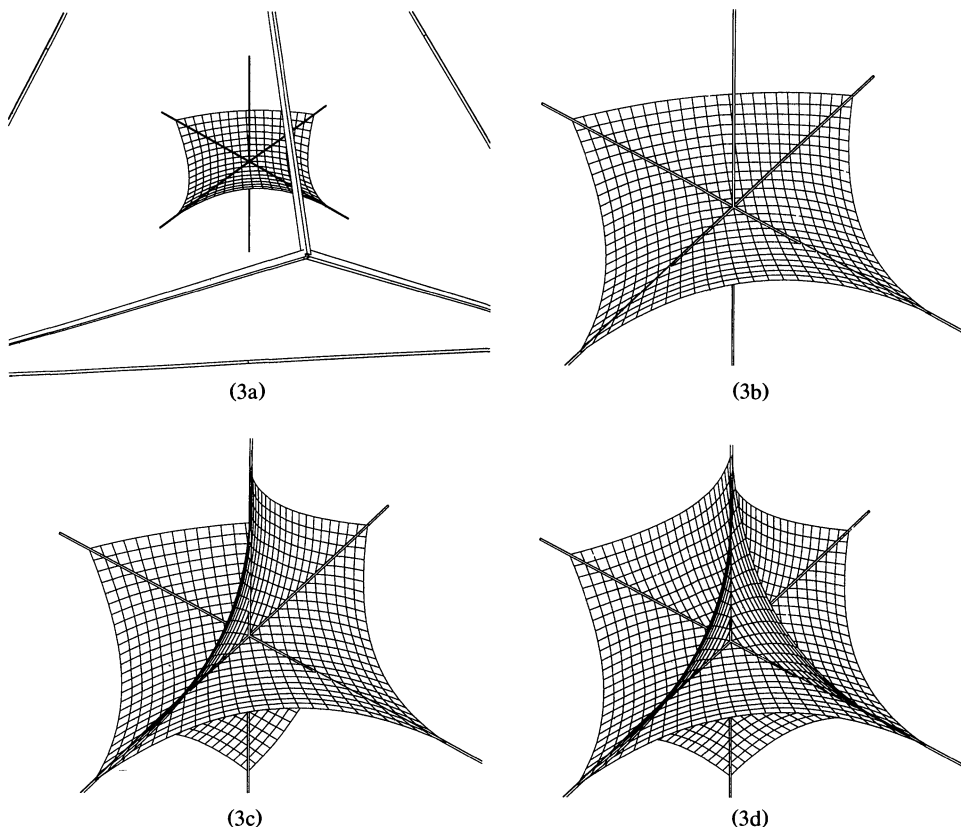


Figure 3. Planes tangent to one of three *saddle-shaped* surfaces bisect the tetrahedron and separate opposing *edges*. The saddles in each pair are tangent at the ends of the line-segment along which they intersect; but at the segment's midpoint (the tetrahedron's centroid) they are orthogonal—this last, for the regular tetrahedron illustrated.

then its associated barycentric coordinates (with respect to V_0, \dots, V_3) will be

$$b_1 = x, \quad b_2 = y, \quad b_3 = z, \quad \text{and} \quad b_0 = 1 - (b_1 + b_2 + b_3).$$

For example: according to (4.1), that portion of E consisting of the envelope of the bisecting planes that separate the face \overline{ijk} from the origin is the surface

$$xyz - 1/54 = 0 \quad \text{with } 1/6 \leq x, y, z \leq 1/3; \quad (5.1)$$

and three other portions of E are found by replacing $x = b_1$, $y = b_2$, and $z = b_3$ here with the three other groups of the four things $\{b_0, \dots, b_3\}$ taken three at a time.

So it remains to obtain a more analytic description of, say, the envelope of the bisecting planes that separate the edge $\overline{0k}$ from the edge \overline{ij} ; as the remaining two portions of E can then be found by substituting each of the two remaining groups of variables associated with each of the other two pairs of edges.

For this we recall that the envelope of a two-parameter family of surfaces

$$g(x, y, z; p, q) = 0$$

can be constructed by requiring that at the same time $g_p (= \partial g / \partial p) = 0$ and $g_q = 0$; thereby determining the envelope in the form (say) of $x = x(p, q)$, $y =$

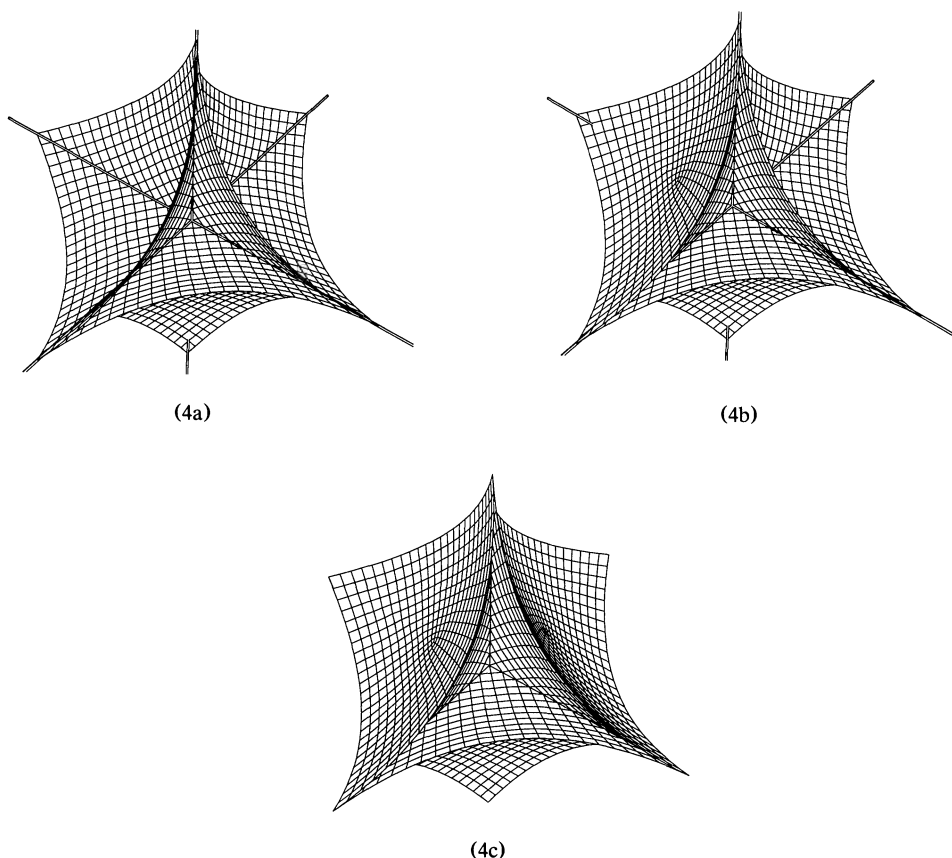


Figure 4. The three saddles fit inside the four cups, thus completing the bisecting envelope of a tetrahedron. Each cup is a cubic. Although each saddle is algebraic, its degree is unknown. If the saddles are regarded as *not* intersecting, then the complete envelope (4c) has the topology of Steiner's *Roman surface*—i.e., of Hilbert and Cohn-Vossen's *heptahedron*. And if part of being an envelope is that the normal line change *continuously*, then the saddles *should* be regarded as not intersecting except at their corners.

$y(p, q)$, and $z = z(p, q)$. We shall do this in the slightly different situation of the three-parameter family of surfaces

$$G(x, y, z; p, q, r) := xp + yq + zr - 1 = 0 \quad (5.2)$$

constrained by a known (albeit yet to be described) function

$$F(p, q, r) = 0 \quad \text{determining } r = r(p, q). \quad (5.3)$$

Here one may verify—with a solution of this last in hand—that the functions $x(p, q)$, $y(p, q)$ and $z(p, q)$ given by

$$x = F_p / (pF_p + qF_q + rF_r), \quad (5.4a)$$

$$y = F_q / (pF_p + qF_q + rF_r), \quad \text{and} \quad (5.4b)$$

$$z = F_r / (pF_p + qF_q + rF_r), \quad (5.4c)$$

indeed solve $0 = g = g_p = g_q$ if we define

$$g(x, y, z; p, q) := G(x, y, z; p, q, r(p, q))$$

as specified in (5.2). The relations (5.4) come about as follows: Eliminating r_p between the p -derivative of (5.2) and that of (5.3) yields

$$x = zF_p/F_r; \quad \text{and so, similarly, } y = zF_q/F_r. \quad (5.5)$$

Substituting (5.5) into (5.2) yields the third relation in (5.4); the first two then follow from (5.5). Favard [9, p. 186] develops relations equivalent to (5.4) in the context of (5.2)–(5.3), albeit under the assumption that F is homogeneous in variables related to those in (5.2).

It remains to construct a function F so that (5.3) implies that the envelope of (5.2) is the envelope of the bisecting planes separating $\overline{0k}$ from \overline{ij} . For this it is more convenient to consider the three coordinate-axis intercepts

$$u := 1/p, \quad v := 1/q, \quad w := 1/r \quad (5.6)$$

of the plane (5.2) regarded—for each fixed u , v , and w —as a surface P in xyz -space. This plane P is to separate the edges mentioned, so we take

$$0 < u, v < 1, \quad \text{and} \quad 1 < w < \infty.$$

The volume of the original tetrahedron U is $1/6$, while the volume of the tetrahedron T bounded by P and the three coordinate planes is $uvw/6$. The relationship $F = 0$ (5.3) is to express the bisection requirement that the amount of T inside U be $1/12$. This quantity, $\text{vol}(T \cap U)$, is computed by finding the intersection of P with the edge \overline{ik} and with the edge \overline{jk} , and subtracting from T 's volume the volume of the tetrahedron inside T but outside U using a sixth of the appropriate scalar triple product. Applying (5.6), the associated F we used in (5.3) was

$$F = r^2 + [p + q - 6 + 2/(pq)]r + 6 - pq - 2(p + q)/(pq), \quad (5.7)$$

$$1 < p, q < \infty, \quad (5.8)$$

and such that

$$0 < r < 1. \quad (5.9)$$

That F here, for p and q given, is quadratic in r facilitates the determination of $r(p, q)$ in (5.3) and the associated surface (5.4). Condition (5.9) is actually a condition on p and q . This actually determines only half of the barycentric coordinates of this sheet of the complete envelope E . The remaining half of the sheet is given by interchanging b_3 ($= z$) and b_0 (this is most easily seen by considering the regular tetrahedron instead of U —a context in which the saddle-shaped character of this sheet is also most apparent). Finally, two other sheets are similarly associated with the other two pairs of edges.

6. THE DEGREES OF THE POLYNOMIALS THAT DESCRIBE THE ENVELOPE. To review: the envelope of the bisecting planes of the tetrahedron has seven parts: three saddle-shaped surfaces and four cup-shaped surfaces. Each of the three saddle-shaped surfaces is associated with separating two edges of the tetrahedron. Each of the four cup-shaped surfaces similarly divides a vertex from a face. We have given the polynomials for the cup-shaped surfaces in (5.1) and its following three lines—they are of total degree three.

We now discuss the polynomials that define the three saddle-shaped surfaces. These polynomials seem to be complicated and we have not been able to complete

this project. The *modus operandi* to find one of these polynomials is to begin with four algebraic equations from §5 in the variables x , y , and z and the parameters p , q , and r . Then one uses resultant theory (see Uspensky [19, Chapter XII]) to eliminate the parameters one by one and terminate with one algebraic equation in the variables x , y , and z . This elimination process was attempted on an 8650 VAX computer using the MACSYMA symbolic manipulation system—but it did not complete the final elimination in many days of standby computer time.

More specifically, using (5.3) and (5.4), we obtain:

$$(pF_p + qF_q + rF_r)x - F_p = 0, \tag{6.1a}$$

$$(pF_p + qF_q + rF_r)y - F_q = 0, \tag{6.1b}$$

$$(pF_p + qF_q + rF_r)z - F_r = 0, \tag{6.1c}$$

$$F = 0. \tag{6.1d}$$

We then substitute into the equations (6.1) the expression for F given by (5.7) and clear fractions to obtain a set of four polynomial equations in the variables x , y , and z and parameters p , q , and r . The parameters are subject to the conditions in (5.8) and (5.9). Again, the condition (5.9) is actually a restriction on p and q . The conditions on p and q have no relevance to the resultant algorithm, which treats the parameters to be eliminated as formal symbols. The restrictions (5.8) and (5.9) arise geometrically. However, they are also algebraic restrictions. We know, for example, that if $r = p = q = 1$, then $F = F_p = F_q = F_r = 0$ and the four equations (6.1a–d) are satisfied regardless of the values of x , y , z . Thus all points in R^3 would be on the surface defined by (6.1) for the values $r = p = q = 1$.

To eliminate p , q , and r from (6.1) requires three applications of the resultant calculation. We had enough machine time to do two of the three. We outline these two using the tables below.

The left side of each equation in (6.1) is multiplied by its denominators' least common multiple, yielding polynomials $f^{(1)}(x, p, q, r)$, $f^{(2)}(y, p, q, r)$, $f^{(3)}(z, p, q, r)$, and $f^{(4)}(p, q, r)$. Their degrees and number of terms are displayed in Table 1. This is the first set of equations to which we apply the resultant operation.

TABLE 1

Equation	Degree	# of terms
$f^{(1)}(x, p, q, r) = 0$	6	12
$f^{(2)}(y, p, q, r) = 0$	6	12
$f^{(3)}(z, p, q, r) = 0$	6	13
$f^{(4)}(p, q, r) = 0$	4	9

Eliminating r from the appropriate pairs of equations in Table 1 yielded Table 2.

TABLE 2

Equation	Degree	# of terms
$g^{(1)}(x, p, q) = 0$	9	21
$g^{(2)}(y, p, q) = 0$	9	21
$g^{(3)}(z, p, q) = 0$	10	39

Eliminating q from the appropriate pairs of equations in Table 2 yielded Table 3.

TABLE 3

Equation	Degree	# of terms
$h^{(1)}(x, y, p) = 0$	32	≈ 500
$h^{(2)}(x, z, p) = 0$	27	≈ 300

Assuming little or no cancellation in the elimination of p from the two equations in Table 3 to obtain a single polynomial equation $P(x, y, z) = 0$, we estimate the degree of P to be 150—and if P has no factors, this seems a surprisingly large degree to be associated with such a simple problem. Unfortunately, the MACSYMA computation attempting this elimination was terminated (without output) after many days of calculation.

On the other hand, C. de Boor felt he could demonstrate a lower bound on the saddle's degree [4]. For this he took $N \approx 45$ points on the horizontal saddle associated with the tetrahedron having as vertices the points $(\pm(3, 3), 3)$ and $(\pm(3, -3), -3)$ —the resulting saddle has *its* corners at $\pm \mathbf{i}$ and $\pm \mathbf{j}$. He applied a numerical algorithm, based on [5], to construct a space \mathcal{Q} of polynomials in three variables of smallest possible degree that allowed unique interpolation to arbitrary values at the N points and at the additional point $(1, -1, 1)$. Although \mathcal{Q} turned out to contain all quartics, he found that nontrivial members of \mathcal{Q} vanishing at the first N points had degree higher than four. This convinced us that the surface defined by (6.1) and (5.7) is of degree higher than four.

7. GRAPHS OF THE ENVELOPE OF THE BISECTING PLANES OF A TETRAHEDRON. For those who care about graphics as well as graphs: The principal software invoked in the computer construction of FIGURES 2–4 was the PLTN2 “super-package,” developed by J. M. Hyman and R. Dougherty to ease (interactively) the application of both M. Prueitt’s GRAFIC package (used here) and the NCAR (National Center for Atmospheric Research) package. All people mentioned in this connection were at the Los Alamos National Laboratory. GRAFIC plots a sequence of surfaces in three dimensions, each surface being prescribed by a “logically rectangular” set of points $(\mathbf{X}_{ij})_{i=1}^m_{j=1}^n$ lying in the surface. The surface is then approximated as follows. The smallest logical sub-squares are each edged by line segments; and (for the purpose of computing normals) the surface spanning these four segments’ is considered to be the two-parameter bilinear average interpolating the four vertices (which is one of the doubly ruled surfaces containing these vertices). GRAFIC removes points and line segments that are hidden from the viewer by other surfaces, and both shading and color are options. Indeed, the most illuminating version of these figures includes both.

For the figures, the edges of the tetrahedron and the axes through its centroid are specified to be slender tubes (with polygonal sections), not lines. The requirement of logically rectangular data meant that each cup-shaped segment of the envelope is graphed as the union of three four-edged sections. Along the two (adjoining) outer edges of one of these sections, the mesh is relatively uniform (each such edge is also an edge of one of the saddles). But along the other two edges (where these sections join each other) the mesh diminishes like $r^{4/3}$ towards

the center of the cup-shaped segment. This was done to improve the visual smoothness (associated with the fact that each “cup” is, in fact, analytic), and roughly approximates C. de Boor’s suggestion to use a coordinate system associated with a conformal map of a 90° angle onto a 120° angle for this purpose.

8. DIVIDING TRIANGLES INTO REGIONS OF UNEQUAL SIZE. The original problem in computational hydrodynamics requires information about the lines (or planes) that divide a region into two subsets of prescribed (and not necessarily equal) relative size.

Towards this end we consider for given θ , $0 < \theta \leq 1/2$, the envelope E_θ of the lines that separate a triangle T into polygons of relative area θ and $1 - \theta$. As in §3, this problem and its solution are invariant under nonsingular affine maps. So, as in §3, it is relevant that there are two equilateral hyperbolas of the form $xy = \text{constant}$ such that some of the tangent lines of each cut the right triangle $\overline{0ij}$ into two such regions. More specifically, these hyperbolas satisfy

$$\text{either } 4xy = \theta \quad \text{or} \quad 4xy = 1 - \theta,$$

depending on whether the θ -fraction lies on the 0-side of the tangent line or on its other side. It follows that: *for given θ , $0 < \theta < 1/2$, the envelope E_θ of the lines that divide a triangle T into portions of relative size θ and $1 - \theta$ consists of three pairs of hyperbolic segments. One pair of the three pairs is associated with each vertex V_i of T by having as asymptotes the two edges of T that contain V_i ; and the barycentric coordinates of this pair satisfy (with subscripts taken mod 3) $b_{i-1} + b_i + b_{i+1} = 1$, together with*

$$(a) \quad 4b_{i-1}b_{i+1} = \theta \quad \text{or} \quad (b) \quad 4b_{i-1}b_{i+1} = 1 - \theta. \quad (8.1a)$$

At the ends of each hyperbolic segment (say of type (a) for some $i = i_1$) one switches to another of the other type (i.e., type (b) for some $i \neq i_1$). Hence, the point (b_{i-1}, b_i, b_{i+1}) at the V_{i+1} -end of the type (a)-segment, being also at the V_{i+1} -end of a type (b)-segment, satisfies

$$4b_{i-1}b_{i+1} = \theta, \quad 4b_i b_{i+1} = 1 - \theta, \quad \text{and} \quad b_{i-1} + b_i + b_{i+1} = 1. \quad (8.1b)$$

Consequently: *The endpoints of the hyperbolic segments (8.1a) or (8.1b) also satisfy*

$$b_{i \pm 1} = 1/2 \quad (\text{so that, also, } b_{i \mp 1} + b_i = 1/2 \text{ there}). \quad (8.1c)$$

Checking that, indeed, the two types of hyperbolic segments are tangent at such common points, we see that these cusps (where the interlaced hyperbolic segments now join together with continuously turning tangent line) trace (for $0 < \theta \leq 1/2$) the three open line-segments whose closures connect the three midpoints of the edges of T . And the cusps perform this covering in one-to-one fashion.

This is all illustrated in FIGURE 5. There it is seen, as θ moves from $1/2$ to 0, that the original three-cusped envelope $E_{1/2}$ (FIGURE 1) doubles its length for θ just below $1/2$, becomes a trefoil containing T ’s centroid for $\theta = 4/9$; and that E_θ approaches the boundary of the original triangle T as θ approaches zero. Except for $\theta = 1/2$ or $4/9$, the hyperbolas composing E_θ cross (transversely) thrice (each time on one of T ’s medians), so that it is seen that there is no θ , $0 < \theta \leq 1/2$, such that E_θ bounds only a convex figure.

Let us now illustrate how such envelopes could be used to locally approximate a smooth boundary between a plane region D , colored white, and its black-colored complement, when given as data only the average color $\int_T \chi_D dA / \int_T dA$ of each triangle T in a tessellation of the plane into small, equilateral triangles. (χ_D , here,

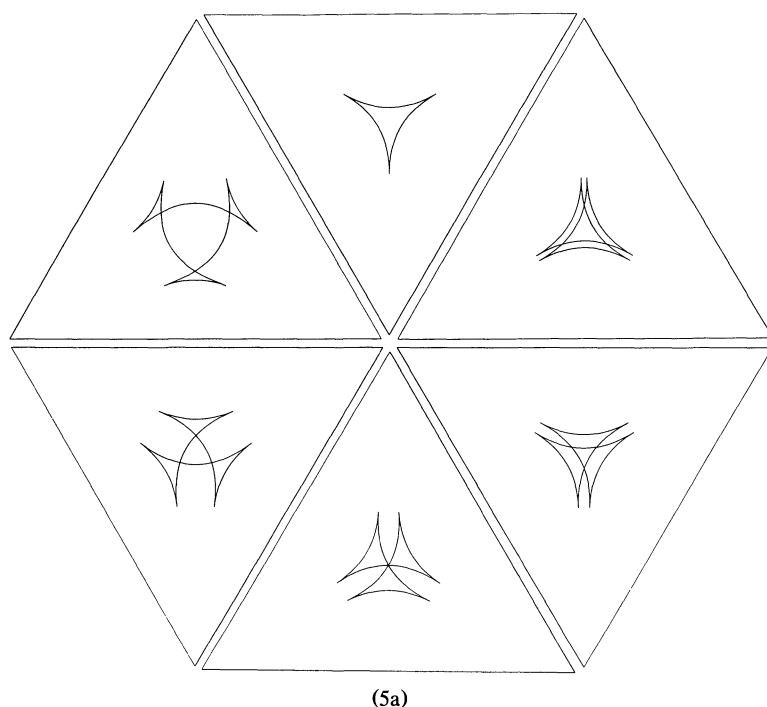
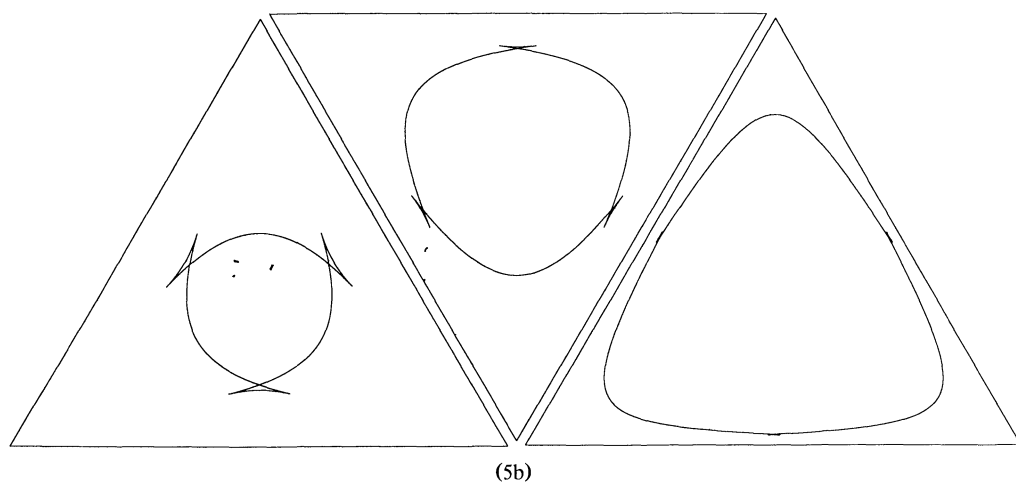


Figure 5. The envelopes of the lines that divide a triangle into two pieces having relative areas θ and $1 - \theta$. Above: clockwise from the top, $\theta = 1/2, 0.485, 0.47, 4/9, 0.4$, and $1/3$. Below: from the left, $\theta = 1/4, 0.15$, and 0.05 . As the text explains, all curves are segments of hyperbolas whose asymptotes are the sides of the triangle, and all cusps lie on one of the lines joining the midpoints of the original triangle's sides.



is the characteristic function of D , and “small” compares the diameter of T to the curvature of the interface). The average color of most triangles will be either black ($= 0$) or white ($= 1$), but those through which the boundary ∂D passes will be colored intermediate values of gray. Since the triangles are small we could hope that a locally linear approximation of ∂D would be second-order accurate (i.e., have an error that goes to zero like the area, not just the diameter, of each triangle). And, indeed, this will be so if (a) the algorithm reproduces an arbitrary linear boundary exactly, and (b) the construction is both local (i.e., determined by nearby data and used only nearby) and stable.

For example: Suppose, in FIGURE 5, that the triangle at 6 o'clock had average color $4/9$ and the triangle at 10 o'clock had average color $2/3$, and that we ignore the remaining triangles. Then there exist only a finite number of lines that divide each triangle into two pieces having areas of appropriate relative size—namely, the four common tangents to the indicated envelopes. But, only two of these four lines will do as borders of a half-space to approximate ∂D ; since either of the other two would have to be black on one side in order to color one triangle appropriately, but be white instead on that same side to simultaneously color the other triangle appropriately. Like the sign of a square root, the selection between the two remaining candidates for a linear boundary must be made using an additional criterion—for example, the location of some completely white triangle would usually suffice.

Further details will be found in [17], including discussion of geometric circumstances leading to second-order accuracy (and others, to lower-order accuracy in spite of reproducing linear boundaries). That three gray average colors can determine approximating planes in three dimensions is also noted there, along with connections of these problems with polynomial spline functions and with apparently nontrivial generalizations of the “ham-sandwich” problem of Steinhaus.

9. REMARKS. It is worth noting the connection of our envelopes with “surfaces of flotation” (using the terminology of hydrostatics and of naval architecture). Thus, suppose one is given a body K in Euclidean n -space along with a prescribed θ in $(0, 1)$. Let E_θ be the envelope of those (hyper)planes dividing K in two parts having relative volume θ and $1 - \theta$. (If the body K had specific gravity θ (or $1 - \theta$) and were floating in some orientation, then its sea-level (hyper)plane section would be tangent to E_θ independent of that orientation.) In this regard, then, White and John [20] claim to be the first to recognize (1871) that for two dimensions the complete curve of flotation of an object can contain cusps—indeed, our FIGURE 5 for $\theta = 1/4$ is qualitatively described there quite accurately [20, p. 93]. In fact, the curves of flotation (for triangles of unspecified specific gravity) in FIGURES 3, 4 and 5 of their Plate V (kindly sent us by the Secretary of the Royal Institution of Naval Architects) are in harmony with the envelopes E_θ in our FIGURE 5 for $\theta = 4/9$, $1/4$, and $1/3$, respectively. Moreover, the curve of flotation for actual vessels—see, e.g., [20, FIGURE 1 of Plate IV] or the reproduction (from another paper by White) in Greenhill [11, p. 160]—have many characteristics in common with the curves in our FIGURE 5.

In the expanded version of this paper (the report [2]) we included an attempt to use a classic result concerning surfaces of flotation to help demonstrate that the topology of bisecting hypersurfaces $E_{1/2}$ in n -space is that of the projective hyperplane P^{n-1} , but that the topology E_θ for $\theta \neq 1/2$ is, instead, that of the surface S^{n-1} of the unit ball. The attempt fails—but it is both amusing and instructive.

It is possible for the envelope of the bisecting hyperplanes of a region to be a single point—consider a rectangle or a circular disc. When it exists we have called such a point a halfway point (all this in another report [3]). There we extend the concept—i.e., of the notion of the median of a distribution—as follows. Let ρ be a nonnegative function on R^n whose integral R^n is finite. Then a point h in R^n is called the *halfway point* for ρ if any $(n - 1)$ -dimensional hyperplane H containing h has half the mass of ρ on each side: i.e.

$$\int_{H^+} \rho(x) dx = \int_{H^-} \rho(x) dx,$$

where H^+ and H^- denote the two half spaces on either side of H . In [3] we consider some characteristics of functions ρ that have halfway points.

ACKNOWLEDGMENTS. The authors wish to thank the following people for fruitful discussions and help: C. de Boer, E. Calabi, G. D. Chakerian, R. Dougherty, R. Hersch, H. B. Lawson, R. D. Mauldin, and J. Mycielski. The U.S. Department of Energy supported our work.

APPENDIX. FINISHING THE PROOF OF THE PROPOSITION IN §4. Let e_j be the j th coordinate unit vector in the canonical basis for R^n . Note that a point $\xi = \sum_{j=1}^n \xi_j e_j$ in R^n then has *barycentric* coordinates (b_0, b_1, \dots, b_n) relative to the $n + 1$ vectors e_1, \dots, e_n , and the origin $0 =: e_0$ that are given by $b_1 = \xi_1, \dots, b_n = \xi_n$, and $b_0 := 1 - \sum_{j=1}^n b_j$; for then $\xi = \sum_{j=0}^n b_j e_j$ with $\sum_{j=0}^n b_j = 1$. With this the tangent hyperplane at a point $b = \sum_{j=1}^n b_j e_j$ in the hypersurface G of points $b > 0$ satisfying $g(b) := \prod_{j=1}^n b_j - 1/(2n^n) = 0$ (see (4.1) with $i = 0$), namely the hyperplane H_b of points $B = \sum_{j=1}^n B_j e_j$ satisfying $\sum_{j=1}^n (B_j/b_j) = n$, defines (with the n coordinate hyperplanes) a simplex Σ_b whose volume is half that of the standard simplex $\Sigma := \overline{e_0 e_1 \dots e_n}$ (the overbar here means “convex hull of”). (To see that g is indeed an appropriately scaled version of the function f in the discussion above the Proposition in §4—and thus that the equation in (4.1) is correct—note that the simplex S_2 being bisected there is the convex hull of the origin $0 =: w_0$ and the n vectors $w_j := n2^{1/n} e_j$, $1 \leq j \leq n$; so that for x there to satisfy both $x = \sum_{j=1}^n x_j e_j$ and its barycentric expression $x = \sum_{j=0}^n b_j w_j$ relative to w_0, \dots, w_n it suffices that $g(b) = 0$ and $\sum_{j=0}^n b_j = 1$.)

The upper bounds $1/n$ in the inequalities $1/(2n) \leq b_j \leq 1/n$ in (4.1) are simply the additional constraints on b in G appropriate for H_b to bisect Σ ; i.e., for Σ_b to be completely contained in Σ ; i.e., for each of H_b 's n coordinate-axis intercepts nb_j to lie between 0 and 1. The lower bounds $1/(2n)$ on all but b_0 come about as follows: Fix k , $1 \leq k \leq n$. Then b_k satisfying $g(b) = 0$ attains its minimum relevant value $b_k = 1/(2n)$ when the b_j , $1 \leq j \leq n$ but $j \neq k$, all take on their maximum relevant values $1/n$ —and note, for this b in G , that b_0 is also $1/(2n)$.

Finally, we now shall see that b_0 can be no smaller than $1/(2n)$ for all b in the “bisecting subset” $G_{1/2} \subset G$ (i.e., when $\Sigma_b \subset \Sigma$). Equivalently, we shall show that $L(b) := \sum_{j=1}^n b_j$ is maximized on $G_{1/2}$ when $g(b) = 0$ (of course) and all but one of b_1, \dots, b_n are $1/n$. First: as $\nabla L = \sum_{j=1}^n e_j$, L has only one extreme value on the hypersurface G —namely, when all nb_j are $1/(n2^{1/n})$ —and it is a minimum (namely, $2^{-1/n}$). Consequently, maxima for L over $G_{1/2}$ occur on its boundary. But b lies in the boundary of $G_{1/2}$ if and only if at least one b_j is $1/n$, i.e., at least one of the intercepts nb_j of the corresponding bisecting hyperplane H_b is 1. (For if all nb_j are strictly between 0 and 1—and also restricted by $g(b) = 0$ of course—then b is in the interior of $G_{1/2}$ in that the intercepts of H_b then have $n - 1$ independent degrees of freedom locally.) So, suppose $b_n = 1/n$. Then we

wish to maximize $L_{n-1}(b_1, \dots, b_{n-1}) := \sum_{j=1}^{n-1} b_j$ subject to $g_{n-1}(b_1, \dots, b_{n-1}) := \prod_{j=1}^{n-1} b_j - 1/(2n^{n-1}) = 0$. The one interior extremum is again a minimum (with all $n-1$ variables now $1/(n2^{1/(n-1)})$); the boundary containing the maxima again consists of vectors with at least one more intercept of H_b being 1, i.e., with one more coordinate, say b_{n-1} , being fixed at $1/n$. And so forth, down to the point when $L_2(b_1, b_2) := b_1 + b_2$ is to be maximized subject to $g_2(b_1, b_2) := b_1 b_2 - 1/(2n^2) = 0$. Hence (say) $b_2 = 1/n$ and $b_1 = 1/(2n)$. And we have shown what we desired—namely, that the maxima of L over $G_{1/2}$ (and hence the minima $1/(2n)$ of b_0) occur with $b_j = 1/(2n)$ for some $1 \leq j \leq n$, and all the rest at $1/n$.

Geometrically, these extrema occur when the hyperplane H_b that bisects Σ also contains one of the $(n-2)$ -dimensional “edges” of the “face” $\overline{e_1 e_2 \dots e_n}$ of Σ that H_b is separating from the vertex e_0 —each of these n “edges” consists of the convex hull of all but one of the basis vectors e_1, \dots, e_n .

The minimization of b_0 above is a simple example of solving a problem in *geometric programming*, that is, finding the extreme values of a generalized polynomial in n variables subject to generalized polynomial constraints. A generalized polynomial, here, is a linear combination of products of (not necessarily integral) powers of the variables.

Added in Proof. Our text associates the idea of the heptahedron with the names of Hilbert and Cohn-Vossen. However, François Apréy’s recent and handsome book, *Models of the Real Projective Plane* (Friedr. Vieweg & Sohn, Braunschweig, 1987), calls it (p. 17) the *Reinhardt heptahedron*. An appropriate reference is: Curt Reinhardt, *Zu Möbius’ Polyedertheorie*, *Berichte über die Verhandlungen der Königlichen Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-physikalische Classe*, vol. 37, 1885, pp. 106–125. Reinhardt also deposited a cardboard model in the Mathematical Institute of the University at Leipzig. We do not know if it survived.

REFERENCES

1. P. Appell, *Équilibre et mouvement des milieux continus*, *Traité de mécanique rationnelle*, vol. 3, 3rd edition, Gauthier-Villars, Paris, 1921.
2. W. A. Beyer and B. Swartz, *The envelope of the planes that bisect a tetrahedron*, Los Alamos report LA-UR-90-2491, (July, 1990).
3. W. A. Beyer and B. Swartz, *Halfway points*, *SIAM J. Math. Anal.*, 23 (1992), 1332–1341.
4. C. de Boor, private communication, 1990.
5. C. de Boor and A. Ron, *On multivariate polynomial interpolation*, *Constr. Approx.* 6 (1990), 287–302.
6. H. Bouasse, *Hydrostatique: Manomètres, baromètres, pompes; Équilibre des corps flottants*, Librairie Delagrave, Paris, 1923.
7. R. Courant, *Differential and Integral Calculus*, vol. 2, Interscience, New York, 1961.
8. R. Courant and F. John, *Introduction to Calculus and Analysis*, vol. 2, Wiley-Interscience, New York, 1974.
9. J. Favard, *Cours de géométrie différentielle locale*, Gauthier-Villars, Paris, 1957.
10. G. K. Francis, *A Topological Picturebook*, Springer-Verlag, New York, 1987.
11. A. G. Greenhill, *A Treatise on Hydrostatics*, Macmillan, London, 1894.
12. D. Hilbert and S. Cohn-Vossen, *Geometry and the Imagination*, Chelsea, New York, 1952, (Translation of “Anschauliche Geometrie,” J. Springer, Berlin, 1932, which was reprinted by Dover in 1944).
13. R. H. Jones, *Folding polyhedra*, *Structural Topology* 7 (1982), 45–50.
14. H. Lamb, *Statics: Including Hydrostatics and the Elements of the Theory of Elasticity*, 3rd edition, Cambridge Univ. Press, Cambridge, 1928.
15. M. Spivak, *A Comprehensive Introduction to Differential Geometry*, vol. 1, 2nd edition, Publish or Perish, Berkeley, CA, 1979.

16. D. J. Struik, *Differential Geometry*, Addison-Wesley, Cambridge, MA, 1950.
17. B. Swartz, *The second-order sharpening of blurred smooth borders*, *Mathematics of Computation* 52 (1989), 675–714.
18. G. B. Thomas, *Calculus and Analytic Geometry*, 3rd edition, Addison-Wesley Co., Reading, MA, 1960.
19. J. V. Uspensky, *Theory of Equations*, McGraw-Hill, New York, 1948.
20. W. H. White and W. John, *On the calculation of the stability of ships, and some matters of interest connected therewith*, *Trans. Royal Institution of Naval Architects* (1871), 77–127, (esp. Plates IV, V, and VI—but, “The drawings, not published with the paper, are in the possession of the Institution, and can at any time be seen by members in the Library.—ED.”, according to *Trans. I. N. A.*, 1884, p. 74).

Los Alamos National Laboratory

T Division, MS-B284

Los Alamos, NM 87545

beyer@lanl.gov; bks@lanl.gov

Postscript. We’ve just been pointed (by L. M. Kelly) to G. Gunther and J. B. Wilker’s paper *The bisectrix of a tetrahedron*, *Mathematika* 39 (1992), 93–103. Although figures and any historical perspective are lacking there, we do want to otherwise note a number of similar ideas.

The name of Professor FELIX KLEIN, of the University of Göttingen, together with those of six other German educators, has been cancelled from the roll of honorary members of the National Education Association in response to a persistent demand from active members of the association, from members of the Council of National Defense, and from others.

: ‘—*American Mathematical Monthly*
25, (1918) p. 331

More on Rectangles Tiled by Rectangles

D. G. Mead and S. K. Stein

The first theorem on a rectangle tiled by rectangles was proved by Dehn in 1903 [3]:

Theorem 1. *Let R be a rectangle that has at least one edge of rational length. Let R be tiled by smaller rectangles each of which has the property that the ratio of its length to its width is rational. Then all the edges of R and of the tiling rectangles have rational lengths.*

In 1940 Brooks, et al. [2] obtained this result by associating an electrical network consisting of currents, voltages, and resistances with the tiling and using well known properties of such networks. We will modify their approach slightly by adding a battery to each edge to obtain the following theorems.

Theorem 2. *Let R be a rectangle that has at least one edge of rational length. Let R be tiled by smaller rectangles each of which has a rational perimeter. Then all the edges of R and of the tiling rectangles have rational lengths.*

If the assumption on R in Theorem 2 is replaced by “ R has a rational perimeter,” the result is false, as is shown by tiling the $2\sqrt{2}$ by $4 - 2\sqrt{2}$ rectangle by four $\sqrt{2}$ by $2 - \sqrt{2}$ rectangles.

Theorem 3. *Let R be a rectangle that has at least one edge of rational length. Let R be tiled by smaller rectangles whose length and width differ by a rational number. Then all the edges of R and of the tiling rectangles have rational lengths.*

Note that either Theorem 1 or Theorem 3 implies that in a tiling of a rectangle with a rational edge by squares, all the dimensions of the rectangles are rational.

Theorem 4. *Let R be a rectangle whose width and length differ by a rational number. Let R be tiled by smaller rectangles each of which has a rational perimeter. Then all the edges of R and of the tiling rectangles have rational lengths.*

1. THE METHOD. Let G be a connected linear graph with m vertices and n edges, e_1, e_2, \dots, e_n , such that each edge is incident to two distinct vertices. There may be more than one edge incident to the same vertices. If edge e_i is incident to the vertices A and B we orient e_i by selecting one of the orientations AB or BA . (We may think of the orientation AB as an arrow from A to B and the algebraic boundary of e_i as $B - A$).

Let C_1 consist of the formal sums $\sum_{i=1}^n x_i e_i$, where x_i is real. Such a sum is shorthand for a function h from the set of edges to the real numbers, where

$h(e_i) = x_i$. C_1 is a vector space of dimension n with real coefficients. If the vertices are V_1, V_2, \dots, V_m let C_0 consist of the formal sums $\sum_{i=1}^m y_i V_i$ where the y_i are real. This sum stands for a function p from the set of vertices to the real numbers, where $p(V_i) = y_i$. Define $\partial: C_1 \rightarrow C_0$ by setting $\partial(e_i) = B_i - A_i$ if e_i is oriented from A_i to B_i and extending by linearity. Note that $p(\partial e_i) = p(B_i - A_i) = p(B_i) - p(A_i)$.

Let r_1, r_2, \dots, r_n be nonnegative real numbers associated with the edges e_1, e_2, \dots, e_n respectively such that the set of edges associated with the r_i 's that are 0 contains no closed circuit. Let w_1, w_2, \dots, w_n be n real numbers.

Consider the following two equations for the unknown functions $h \in C_1$ and $p \in C_0$:

$$\text{I.} \quad \partial \left(\sum_{i=1}^n h(e_i) e_i \right) = 0$$

$$\text{II.} \quad p(\partial e_i) = w_i - r_i h(e_i).$$

(In terms of electrical networks, $h(e_i)$ is the current in e_i , r_i is the resistance in e_i , w_i is the electromagnetic force of a battery attached to e_i and $p(V)$ is the potential at V . Equation I asserts that the total current entering a vertex is 0. Equation II relates the voltage drop over an edge to the current, resistance and the strength of the battery at that edge.)

In ([1], 162–171) it is shown that these simultaneous equations have a unique solution for h and a unique, up to an additive constant, solution for p . Actually, in [1] it is assumed that all r_i are positive. However, the key argument, which appears in the footnote on p. 171, goes through with our weaker assumptions, as long as the spanning tree used in the proof is chosen to contain the edges for which $r_i = 0$. (There is a misprint in the footnote: the final \geq should be replaced by $>$.) Moreover, the formulas for the values of h and p obtained there show that if r_i and w_i , $1 \leq i \leq n$, are all rational, then so are the values of h and therefore of $p(\partial e_i)$. The same conclusion holds if “nonnegative” is replaced by “nonpositive” in [1].

As in [2] associate a linear graph with a tiling of a rectangle R by rectangles R_1, R_2, \dots, R_{n-1} . (For convenience, denote R also by R_n .) To do this, introduce an xy coordinate system such that R_n is in the first quadrant, its edges are parallel to the axes, and the origin is at a corner of R . Each rectangle R_i , $1 \leq i \leq n$, has edges parallel to the x -axis (the “horizontal edges”) of length h_i and edges parallel to the y -axis (the “vertical edges”) of length v_i .

Let S be the union of all the horizontal edges of the n rectangles. The midpoints of the connected components of S will be the vertices of a linear graph G . For each rectangle R_i , $1 \leq i \leq n-1$ introduce an edge oriented from the component containing its lower edge to the component containing its upper edge. For $R_n = R$, introduce an edge oriented from its upper edge down to its lower edge. At a vertex V define $p(V)$ to be the y -coordinate of V . Thus for $1 \leq i \leq n-1$, $p(\partial e_i) = v_i$ and $p(\partial e_n) = -v_n$. Also define $h(e_i)$ to be h_i , $1 \leq i \leq n$.

The definitions of r_i and w_i will depend on the particular theorem to be proved.

2. PROOFS OF THE THEOREMS. The proof of Theorem 1, as given in [2], goes as follows. First place R in such a way that its vertical length v_n is rational. Define r_i to be $-v_i/h_i$, $1 \leq i \leq n-1$. Define r_n to be 0. Define w_i to be 0, $1 \leq i \leq n-1$ and w_n to be $-v_n$. Checking that I and II are satisfied is straightforward. Thus, all h_i and v_i , $1 \leq i \leq n$, are rational.

To prove Theorem 2 let $w_i = h_i + v_i$, $1 \leq i \leq n - 1$, and $w_n = -v_n$. Let $r_i = 1$, $1 \leq i \leq n - 1$, and $r_n = 0$.

To prove Theorem 3 first place the rational edge of R along the y -axis. For $1 \leq i \leq n - 1$ let $r_i = -1$ and $w_i = v_i - h_i$. Let $r_n = 0$ and $w_n = -v_n$.

To prove Theorem 4 let $r_i = 1$ and $w_i = h_i + v_i$ for $1 \leq i \leq n - 1$. Let $r_n = 1$ and $w_n = h_n - v_n$.

These theorems could be generalized by assuming, that for each R_i , $1 \leq i \leq n - 1$, there is a positive (negative) rational number r_i such that $v_i + r_i h_i$ is rational and that v_n is rational. The proof is similar.

3. ANOTHER APPROACH. In the proof in [1] the values of p , h , and w are never multiplied by each other. Thus we may take their values in a vector space over the field generated by r_1, \dots, r_n , in particular in the abelian group \mathbb{R}/\mathbb{Q} , under addition. In the proofs w_i is now an element of \mathbb{R}/\mathbb{Q} , the zero element. Equations I and II, which refer to elements in \mathbb{R}/\mathbb{Q} , hold and again the solution is unique, namely p and h must both be the constant function with value $0 \in \mathbb{R}/\mathbb{Q}$.

For a different type of problem concerning tiling a rectangle by rectangles see [4].

REFERENCES

1. D. W. Blackett, *Elementary Topology*, Academic Press, New York, 1967.
2. R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte, The dissection of rectangles into squares, *Duke Math. J.* 7 (1940), 312–340.
3. M. Dehn, Zerlegung von Rechtecke in Rechtecken, *Math. Ann.* 57 (1903), 314–332.
4. S. Wagon, Fourteen proofs of a result about tiling rectangles, *Amer. Math. Monthly* 94 (1987), 601–617.

Mathematics Department
University of California at Davis
Davis, CA 95616-8633

The Birthday Problem

In an article in the January, 1992, issue of the MONTHLY, Joag-Dev and Proschan present an elementary example of the use of majorization in probability. This example considers the Birthday Problem where different dates have different probabilities.

Another frequently taught problem in probability is the Coupon Collector's Problem, and this problem provides a similar elementary example of the use of majorization. Suppose that n objects are picked repeatedly and independently with the probability that object i is picked at on a given try is p_i (where $p_1 + \dots + p_n = 1$). Let $\mathbf{p} = (p_1, \dots, p_n)$ and let $T_{\mathbf{p}}$ be the Coupon Collector's Time, i.e. the earliest time where all n objects have been picked at least once. A reasonable exercise for someone who has read the article of Joag-Dev and Proschan is to show that

$$P(T_{(1/n, \dots, 1/n)} \leq t) \geq P(T_{\mathbf{p}} \leq t)$$

for any \mathbf{p} and to show that $P(T_{\mathbf{p}} \leq t)$ is a Schur-concave function of \mathbf{p} .

—Martin V. Hildebrand
Department of Mathematics
The University of Michigan
Ann Arbor, MI 48109-1003

Ramanujan—For Lowbrows

Bruce C. Berndt and S. Bhargava

“No, Inspector,” he said. “It is not at all like that, I am assuring you. You see, for a person of my sort—and I admit that we are a rare breed—numbers are so much in our minds there is hardly any question of writing them down, let alone adding one to another.” . . .

“Let me give you one instance,” he said. “Before I was beginning work just now, I was taking a short stroll, and I happened to see a handcartwalla. Now, being the sort of chap I am, I of course notice the number burned on the side of the cart: seventeen-twenty-nine. Now, does that mean anything to you yourself?”

“It is the number on the cart,” Ghote answered guardedly. “By law it must be there.”

Raghu Barde smiled his warm smile again.

“Ah, yes, the police view. But what do you think those figures meant to me? You would never guess. But the moment I was seeing them I said: Aha, the smallest number expressible as a sum of two cubes in two different ways. And, you know, if ever I am getting to marry, I suppose I will want a wife whose birth date comes to some number pleasing to me like that.”

“I see,” Ghote said.

And, although the mumbo jumbo about cubes and expressible meant nothing to him, and he could not help thinking that to choose a wife by number would be a much riskier proceeding than to let the astrologers choose one for you, he did dimly see what a different sort of life Raghu Barde lived from that of the common number-unencumbered man.

H. R. F. Keating
Dead on Time

1. INTRODUCTION. To celebrate the centenary of Ramanujan’s birth, in June, 1987, an international conference was held at The University of Illinois at Urbana-Champaign [1]. Numerous roads through varied scenery brought researchers from Ramanujan’s papers, problems, letters, notebooks, and unpublished manuscripts to a panoply of areas of contemporary research, including partitions, mock theta-functions, statistical mechanics, Lie algebras, probabilistic number theory, modular forms, elliptic functions, complex multiplication, hypergeometric series, q -series, asymptotic expansions, and beta integrals. Very few mathematicians have ever had such a broad impact on mathematical research. Although many results presented at the conference could be understood and appreciated by mathematicians outside these areas of research, this was a conference for *highbrows*.

Many of Ramanujan’s beautiful discoveries, however, are easily understood, are elementary, and appeal to a wide variety of tastes. Thus, this paper is written for *lowbrows*. Only elementary algebra is needed to prove the lion’s share of theorems reported here. Most are found in the unorganized portion of Ramanujan’s second notebook, his third notebook, and problems that he posed for readers of the *Journal of the Indian Mathematical Society*. The results we describe fall under the headings of elementary algebra, equal sums of powers, and elementary number theory.

We begin our expedition in a taxi-cab as we recount G. H. Hardy's riding in taxi-cab no. 1729 to visit Ramanujan while lying ill in Putney. Some historical remarks are offered on the two representations $1^3 + 12^3 = 9^3 + 10^3$ of 1729. This leads us to Euler's solution, rediscovered by Ramanujan in a simpler form, of the diophantine equation $A^3 + B^3 = C^3 + D^3$.

We turn from equal sums of third powers to equal sums of fourth powers and ask "Did Ramanujan ever read *Mathematical Magazine*?" No, we are not speaking of the journal, *Mathematics Magazine*, published by the MAA, with the first issue appearing under a slightly different title in 1926, six years after Ramanujan's death. Some historical remarks will be made about *Mathematical Magazine*.

We next temporarily stop our journey to view what the authors consider to be one of the most captivating, enthralling finite identities in all of mathematics. Is this marvelous identity simply an accident on the road to sums of powers? Or are we at the base of the Himalayas—facing away from the mountains?

We next encounter three types of systems of equations. The first system leads us to sequences that decrease for a while, then increase for a while, etc. We must have roamed to a college campus, for these sequences involve radicals, infinitely many of them. Like most radicals, these have interesting properties. The second system leads us to a visit with S. Ramanujan. No, that is not a misprint! Is he really Ramanujan, or is he someone else? Our third system was solved beautifully by Ramanujan in his third published paper, but he did not realize that J. J. Sylvester had solved this system in 1851, nor was Ramanujan aware of the implications of his work. We provide a sketch of Ramanujan's clever proof.

Proceeding from a sketch to a complete landscape, we provide proofs of some interesting properties of roots of cubic polynomials that Ramanujan discovered. As applications, we offer two curious trigonometric identities.

For our last proof, we establish sharp bounds for a sum giving the largest power of a prime dividing $n!$.

We conclude our paper with some approximations to π .

Several references will be made to Ramanujan's notebooks [26], published in two volumes. The second volume contains the second and third notebooks, and all page numbers in this paper refer to the pagination in this volume.

2. SUMS OF POWERS. Many readers are familiar with the famous taxi-cab story immortalized by Hardy [27, p. xxxv]. "I remember once going to see him when he was lying ill at Putney. I had ridden in taxi-cab no. 1729, and remarked that the number seemed to me rather a dull one, and that I hoped it was not an unfavourable omen. 'No,' he replied, 'it is a very interesting number; it is the smallest number expressible as a sum of two cubes in two different ways.'" (It is clear that the author of the opening passage about a handcart with 1729 imprinted on its side was acquainted with this delightful incident in the life of Ramanujan and Hardy. A handcartwalla is a person who pulls a two-wheeled handcart, normally carrying one or two people, and is no longer a common sight in present day India. The suffix "walla" comes from Hindi.) In fact, Ramanujan had previously recorded these two representations for 1729, $1^3 + 12^3$ and $9^3 + 10^3$, on page 225 of his second notebook [26]. However, this example appears to have been first noticed by B. Frénicle de Bessy in 1657. Frénicle and J. Wallis each found additional examples for two equal sums of two cubes. A bitter argument ensued with each accusing the other of using trivial methods. Since P. Fermat also frequently was feuding with these two men, letters detailing their acrimony can be

OXFORD EDITION

THE
Poetical Works
OF
WILLIAM WORDSWORTH

WITH INTRODUCTIONS AND NOTES

EDITED BY

THOMAS HUTCHINSON, M.A.



LONDON

HENRY FROWDE

OSFORD UNIVERSITY PRESS WAREHOUSE

AMEN CORNER, E.C.

1893

The frontispiece of a volume of Wordsworth's poetry. The volume was awarded to the young Ramanujan for his "outstanding work in Maths." Such prizes for mathematical contests were common in Ramanujan's hometown, Kumbakonam, and throughout India of the period.

found in Fermat's *Oeuvres* [11, pp. 419–420; 427–457] and E. T. Bell's book [2, Chapter 12], as well as in L. E. Dickson's *History* [8, p. 552]. In 1898, C. Moreau [18] found the ten solutions of $A^3 + B^3 = C^3 + D^3$ with the sums less than 100,000. After 1729, the next largest sum is $4104 = 2^3 + 16^3 = 9^3 + 15^3$.

From another viewpoint, Ramanujan provided Hardy with solutions to the classical diophantine equation

$$A^3 + B^3 + C^3 = D^3. \quad (2.1)$$

L. Euler [10] completely solved (2.1) for positive or negative rational solutions. At three places in his notebooks, Ramanujan addresses the problem of finding solutions of (2.1). In Entry 20(iii) of Chapter 18 and on page 266 in the unorganized portion of his second notebook, Ramanujan provides parametric solutions to (2.1), but they are not as general as Euler's. But near the end of his third notebook [26, p. 387], Ramanujan offers a family of solutions equivalent to Euler's general solution. Both Hardy [13, p. 11] and G. N. Watson [30] discussed one of Ramanujan's less general solutions to (2.1). They had no knowledge of Ramanujan's general solution, because they did not have access to the third notebook. We quote Ramanujan's theorem.

Theorem. *If*

$$\alpha^2 + \alpha\beta + \beta^2 = 3\lambda\gamma^2,$$

then

$$(\alpha + \lambda^2 \gamma)^3 + (\lambda \beta + \gamma)^3 = (\lambda \alpha + \gamma)^3 + (\beta + \lambda^2 \gamma)^3. \quad (2.2)$$

As an example, we recover the two pairs of aforementioned taxi-cab cubes by putting $(\alpha, \beta, \gamma, \lambda) = (3, 0, 1, 3)$ in (2.2).

Although several formulations equivalent to Euler's general solution have been discovered, Ramanujan's formulation (2.2) appears to be the simplest of all. The problem of completely characterizing all positive integral solutions of (2.1) is unsolved.

On the other hand, Euler conjectured that there were no positive integral solutions to

$$A^4 + B^4 + C^4 = D^4.$$

It was not until 1988 that Euler's conjecture was shown to be false by N. D. Elkies [9], who found an infinite class of solutions.

Ramanujan derived several theorems providing infinite families of solutions for equal sums of powers. For example, toward the end of this third notebook [26, p. 384], he writes two parametric solutions for representing a fourth power as a sum of five fourth powers.

Theorem. *If s, t, m , and n are arbitrary, then*

$$\begin{aligned} (8s^2 + 40st - 24t^2)^4 + (6s^2 - 44st - 18t^2)^4 + (14s^2 - 4st - 42t^2)^4 \\ + (9s^2 + 27t^2)^4 + (4s^2 + 12t^2)^4 = (15s^2 + 45t^2)^4 \end{aligned} \quad (2.3)$$

and

$$\begin{aligned} (4m^2 - 12n^2)^4 + (3m^2 + 9n^2)^4 + (2m^2 - 12mn - 6n^2)^4 \\ + (4m^2 + 12n^2)^4 + (2m^2 + 12mn - 6n^2)^4 = (5m^2 + 15n^2)^4. \end{aligned} \quad (2.4)$$

Ramanujan recorded several examples. For instance, if we set $s = 1$ and $t = 0$ in (2.3), we find that

$$4^4 + 6^4 + 8^4 + 9^4 + 14^4 = 15^4.$$

Formula (2.3) is due to C. B. Haldeman [12, pp. 289–290] in 1904. Uncannily, Ramanujan used the same notation and recorded the terms in the same order as Haldeman! Likewise, (2.4) was established by Haldeman [12, p. 289] and slightly later by A. Martin [15, pp. 325–326, 331]. Ramanujan does not use Haldeman's notation in (2.4) but does employ Martin's notation!

Ramanujan recorded his results in notebooks from about 1903 until he departed for England in 1914. The 16 chapters in the first notebook and the 21 chapters in the second evince a progressive maturation from more elementary mathematics to much deeper results. The third notebook, however, contains both very elementary results as well as advanced results. While the latter theorems may have been recorded in Cambridge, the former results were probably recorded early in the period 1903–1914. Since in India Ramanujan did not have access to even the primary mathematical journals of his day, it is extremely unlikely that he could have seen the obscure journal, *Mathematical Magazine*, in which Martin and Haldeman published their results. Thus, the notation in (2.3) and (2.4) being identical with that of Haldeman and Martin, respectively, must be coincidental.

Mathematical Magazine was founded and edited by Martin and was devoted to “elementary mathematics.” Issues of the first volume were published quarterly in

1882–1884 at a cost of 50 cents per issue or one dollar per year. The second and last volume of 12 issues was published over the years 1890–1904, with the last four issues appearing in January, 1895; January, 1896; December, 1898; and January, 1904. The last issue contains four papers, three by Martin and one by Haldeman. In the penultimate issue, under the heading “Editorial Items,” we learn that “Since No. 10 of the Magazine was published, three able contributors have ‘crossed over’ and ‘passed beyond the confines of earth.’” It is likely that an even greater number “crossed over” between the 11th and 12th issues. Possibly due to complaints registered by readers disgruntled over the irregularity at which issues appeared, the price per issue had dropped to 30 cents.

Toward the end of the third notebook [26, p. 386], Ramanujan records one of the most fascinating identities we have ever seen.

Theorem. *Let a, b, c , and d denote any numbers such that $ad = bc$. Then*

$$\begin{aligned} & 64\{(a+b+c)^6 + (b+c+d)^6 - (c+d+a)^6 - (d+a+b)^6 \\ & \quad + (a-d)^6 - (b-c)^6\} \\ & \times \{(a+b+c)^{10} + (b+c+d)^{10} - (c+d+a)^{10} - (d+a+b)^{10} \\ & \quad + (a-d)^{10} - (b-c)^{10}\} \\ & = 45\{(a+b+c)^8 + (b+c+d)^8 - (c+d+a)^8 - (d+a+b)^8 \\ & \quad + (a-d)^8 - (b-c)^8\}^2. \quad (2.5) \end{aligned}$$

The hypothesis $ad = bc$ was omitted by Ramanujan, although it does appear as a hypothesis for some related results on the previous page.

We first transcribe (2.5) into a somewhat more transparent form. For each positive integer m , set

$$\begin{aligned} F_{2m}(a, b, c, d) &= (a+b+c)^{2m} + (b+c+d)^{2m} - (c+d+a)^{2m} \\ &\quad - (d+a+b)^{2m} + (a-d)^{2m} - (b-c)^{2m}. \end{aligned}$$

Put $b = ax$, $c = ay$, and $d = axy$, which does not contravene the hypothesis $ad = bc$. Then it is easy to see that

$$F_{2m}(a, b, c, d) = a^{2m} f_{2m}(x, y),$$

where

$$\begin{aligned} f_{2m}(x, y) &= (1+x+y)^{2m} + (x+y+xy)^{2m} - (y+xy+1)^{2m} \\ &\quad - (xy+1+x)^{2m} + (1-xy)^{2m} - (x-y)^{2m}. \quad (2.6) \end{aligned}$$

Hence, (2.5) can be put in the form

$$64f_6(x, y)f_{10}(x, y) = 45f_8^2(x, y). \quad (2.7)$$

We first employed the computer algebra system *Mathematica* to verify (2.7). Next, using *Mathematica*, we attempted to find other identities like (2.7) involving $f_{2m}(x, y)$ for $m \leq 10$, but we were unsuccessful. We fortunately found a much more informative proof of (2.7) that is not merely a verification via computer algebra [6]. We will not repeat that proof here but instead offer a few additional remarks.

By inspection, we easily see that $x = 0, 1, -1, -2, -1/2$ are zeros of $f_{2m}(x, y)$. By symmetry, $y = 0, 1, -1, -2, -1/2$ are also zeros. Since f_{2m} has degree (at most) $2m$ in each of the variables x and y , it follows that $f_2(x, y) \equiv 0 \equiv f_4(x, y)$. In our original notation, we have therefore proved that, if $ad = bc$, then

$$\begin{aligned} (a + b + c)^n + (b + c + d)^n + (a - d)^n \\ = (c + d + a)^n + (d + a + b)^n + (b - c)^n, \end{aligned} \quad (2.8)$$

where $n = 2$ or 4 . These are the aforementioned results that appear on page 385 of [26]. We have therefore returned to the problem of generating equal sums of biquadrates. Although many results have appeared in the literature yielding two equal sums of three biquadrates [8, pp. 653–657], none appear as simple as Ramanujan's identity (2.8).

Are (2.5) and (2.7) merely accidents, or are they a manifestation of some far deeper theorem?

3. ELEMENTARY ALGEBRA. In courses and texts on beginning calculus, students encounter many monotonic sequences in their study of sequences and series. An inquisitive student may ask for naturally occurring examples of sequences that increase for a while, then decrease for a while, etc. As we shall see, some infinite sequences of nested radicals of Ramanujan provide excellent examples.



Mrs. Ramanujan (S. Janaki Ammal) and W. Narayanan, one of her two adopted sons.

In 1914, Ramanujan [22], [27, pp. 327–329] posed the following problem to readers of the *Journal of the Indian Mathematical Society*: Solve completely

$$x^2 = y + a, \quad y^2 = z + a, \quad \text{and} \quad z^2 = x + a. \quad (3.1)$$

Concomitantly, he asked for the evaluation of three infinite sequences of nested radicals. Toward the end of his second notebook [26, pp. 305–307], Ramanujan recorded further and more general results. It is not difficult to see that x is a root of an octic polynomial. This polynomial can be factored over the quadratic field

$Q(\sqrt{4a-7})$ into one quadratic and two cubic factors. These factors are correctly given by Ramanujan in his solution [22], but the factors given in the solution printed in his *Collected Papers* [27, pp. 327–329] contain four sign errors.

From the equalities (3.1), we find that

$$\begin{aligned} x &= \sqrt{a+y} = \sqrt{a+\sqrt{a+z}} = \sqrt{a+\sqrt{a+\sqrt{a+x}}} \\ &= \sqrt{a+\sqrt{a+\sqrt{a+\sqrt{a+\cdots}}}}. \end{aligned} \tag{3.2}$$

Each square root should be considered two-valued, and so we are led to eight infinite sequences of nested radicals corresponding to the eight roots of our octic polynomial. First, we should determine those values of a for which the infinite radical in (3.2) converges. This is not an easy problem, but each of the eight sequences in (3.2) converges at least for $a \geq 2$ [5, Chapter 22]. As a specific example, let

$$\begin{aligned} a_1 &= \sqrt{a}, & a_2 &= \sqrt{a-\sqrt{a}}, & a_3 &= \sqrt{a-\sqrt{a+\sqrt{a}}}, \\ a_4 &= \sqrt{a-\sqrt{a+\sqrt{a+\sqrt{a}}}}, \dots, \end{aligned}$$

where the sequence of signs $-, +, +, \dots$ appearing in the nested radicals has period 3. A careful analysis shows that

$$a_{6n+1} > a_{6n+2} > a_{6n+3} > a_{6n+4}$$

and

$$a_{6n+4} < a_{6n+5} < a_{6n+6} < a_{6n+7},$$

for each nonnegative integer n . Furthermore,

$$0 < a_4 < a_{10} < \cdots < a_{6n+4} < a_{6n+7} < a_{6n+1} < \cdots < a_7 < a_1 = \sqrt{a}.$$

Thus, $\{a_{6n+1}\}$ and $\{a_{6n+4}\}$ converge. Next, it must be shown that $\{a_{3n+1}\}$ converges and, lastly, that $\{a_n\}$ converges. The details in this analysis are not easy [5, Chapter 22].

If we solve the two cubic equations mentioned above, it is not easy, in general, to identify the roots with the appropriate infinite sequences of radicals. For example,

$$\lim_{n \rightarrow \infty} a_n = \frac{A-1}{6} + \frac{2}{3} \sqrt{4A+A} \sin\left(\frac{1}{3} \arctan \frac{2A+1}{3\sqrt{3}}\right), \tag{3.3}$$

where $A = \sqrt{4a-7}$. We made these identifications by expanding both the algebraically determined roots and the infinite radicals around “ $a = \infty$.” For example, both sides of (3.3) have the asymptotic expansions

$$\sqrt{a} - \frac{1}{2} - \frac{3}{8\sqrt{a}} - \frac{1}{4a} + \cdots,$$

as a tends to ∞ . For particular numerical examples, the proper identifications are easier to make. For instance, if $a = 2$ in (3.3),

$$2 \sin\left(\frac{\pi}{18}\right) = \sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 - \cdots}}}}.$$

Later, Ramanujan [25], [27, p. 332] submitted the similar problem of determining the simultaneous solutions of the system,

$$x^2 = a + y, \quad y^2 = a + z, \quad z^2 = a + u, \quad \text{and} \quad u^2 = a + x,$$

to the *Journal of the Indian Mathematical Society*. Fourteen years elapsed before a solution by G. N. Watson [29] was published, while another solution can be found in [5, Chapter 22]. As above, interesting sequences of nested radicals arise. For example,

$$\frac{1}{2}(2 + \sqrt{5} + \sqrt{15 - 6\sqrt{5}}) = \sqrt{5 + \sqrt{5 + \sqrt{5 - \sqrt{5 + \sqrt{5 + \cdots}}}}},$$

where the infinite sequence of signs $+, +, -, +, \cdots$ has period 4.

The theory of infinite sequences of nested radicals has not been well developed, probably because general theorems are difficult to obtain and convergence is slow. For further examples, theorems, and references to the literature, see [3, pp. 108–112] and [5, Chapter 22].

In the unorganized portions of his notebooks [26] and in the problem sections of the *Journal of the Indian Mathematical Society*, Ramanujan offers other problems on systems of equations. Thus, on page 338 of [26], he asks for the solutions of

$$\frac{x^5 - a}{x^2 - y} = \frac{y^5 - b}{y^2 - x} = 5(xy - 1),$$

where a and b are arbitrary constants. There are 25 pairs (x, y) of solutions. The special case $a = 6$, $b = 9$ appeared as Question 284 [20], [27, pp. 322–323] in the *Journal of the Indian Mathematical Society*. Ramanujan's solution was the only one received, and a similar solution to the more general problem can be found in [5, Chapter 22].

Question 284 was the fourth problem that Ramanujan published in the *Journal of the Indian Mathematical Society*. The first five problems that Ramanujan posed to *Journal* readers were published under the name S. Ramanujam. Ramanujan and Ramanujam are two versions of the same Sanskrit name RAMANUJAH, which means younger brother of Rama.

We mention one further system of equations studied by Ramanujan. On page 338 of his second notebook, Ramanujan asks, in slightly different notation, for the solutions of the system of $2n$ equations,

$$x_1 y_1^{j-1} + x_2 y_2^{j-1} + \cdots + x_n y_n^{j-1} = a_j, \quad 1 \leq j \leq 2n, \quad (3.4)$$

where $x_1, \dots, x_n, y_1, \dots, y_n$ are $2n$ unknowns, and in his short paper [21], [27, pp. 18–19], Ramanujan presents his clever solution, which we briefly indicate.

Ramanujan defines

$$\varphi(\theta) := \sum_{j=1}^n \frac{x_j}{1 - \theta y_j}. \quad (3.5)$$

When $\varphi(\theta)$ is expanded in a power series in θ , it is seen that the coefficient of θ^k is a_{k+1} , $0 \leq k \leq 2n - 1$. On the other hand, $\varphi(\theta)$ has the form

$$\varphi(\theta) = \frac{\sum_{j=0}^{2n-1} A_{j+1} \theta^j}{1 + \sum_{j=1}^n B_j \theta^j}. \quad (3.6)$$

Clearing the denominator in (3.6) and using the aforementioned power series for $\varphi(\theta)$, we can determine first the coefficients B_j , $1 \leq j \leq n$, and secondly the

coefficients A_j , $1 \leq j \leq n$, in terms of a_1, a_2, \dots, a_{2n} by equating coefficients of like powers of θ . Having explicitly determined A_j and B_j , $1 \leq j \leq n$, we substitute these values into (3.6) and once again expand $\varphi(\theta)$ into partial fractions. Comparing the result with (3.5), we determine x_j and y_j , $1 \leq j \leq n$.

It is easy to see that the system (3.4) is equivalent to the single equation

$$\sum_{i=1}^n x_i (y_i s + t)^{2n-1} = \sum_{j=0}^{2n-1} \binom{2n-1}{j} a_{j+1} s^j t^{2n-1-j}.$$

Thus, Ramanujan's query is equivalent to the question: When can a binary $(2n-1) - ic$ form be represented as a sum of n $(2n-1)th$ powers? In 1851, Sylvester [28, pp. 203–216, 265–283] found the following necessary and sufficient conditions for a solution: The system of n equations

$$a_j u_1 + a_{j+1} u_2 + \dots + a_{j+n} u_{n+1} = 0, \quad 1 \leq j \leq n,$$

must have a solution u_1, u_2, \dots, u_{n+1} such that the $n - ic$ form

$$p(w, z) := \sum_{j=0}^n u_{j+1} w^j z^{n-j}$$

can be represented as a product of n distinct linear forms. This is true for a general $2n$ -tuple $(a_1, a_2, \dots, a_{2n})$ in the sense of algebraic geometry. Thus, the numbers y_1, y_2, \dots, y_n are related to the factorization of $p(w, z)$. Sylvester's theorem belongs to the subject of invariant theory, which was developed in the late 19th and early 20th centuries. For a contemporary treatment, but with classical language, see a paper by J. P. S. Kung and G.-C. Rota [14].

We next consider the following theorem of Ramanujan [26, p. 325].

Theorem. Let α , β , and γ denote the roots of the cubic equation

$$x^3 - ax^2 + bx - 1 = 0. \quad (3.7)$$

Then, for a suitable determination of roots,

$$\alpha^{1/3} + \beta^{1/3} + \gamma^{1/3} = (a + 6 + 3t)^{1/3} \quad (3.8)$$

and

$$(\alpha\beta)^{1/3} + (\beta\gamma)^{1/3} + (\gamma\alpha)^{1/3} = (b + 6 + 3t)^{1/3}, \quad (3.9)$$

where

$$t^3 - 3(a + b + 3)t - (ab + 6(a + b) + 9) = 0. \quad (3.10)$$

Since this beautiful elementary theorem is evidently new and since a short proof can be given, we provide one here.

Proof: Noting, from (3.7), that $\alpha\beta\gamma = 1$, let

$$z^3 - \theta z^2 + \varphi z - 1 = 0 \quad (3.11)$$

denote the cubic polynomial with roots $\alpha^{1/3}$, $\beta^{1/3}$, and $\gamma^{1/3}$, chosen so that their product equals 1. Cubing both sides of the equality

$$z^3 - 1 = \theta z^2 - \varphi z,$$

we find that

$$(z^3 - 1)^3 - \theta^3 z^6 + \varphi^3 z^3 + 3\theta\varphi z^3(z^3 - 1) = 0. \quad (3.12)$$

Since $\alpha^{1/3}$, $\beta^{1/3}$, and $\gamma^{1/3}$ are roots of (3.11), they are also roots of (3.12). As a cubic polynomial in z^3 , (3.12) thus has the roots α , β , and γ .

Comparing (3.7) and (3.12), we deduce that

$$a = \theta^3 + 3 - 3\theta\varphi \quad (3.13)$$

and

$$b = \varphi^3 + 3 - 3\theta\varphi. \quad (3.14)$$

If we define t by

$$\theta^3 = a + 6 + 3t, \quad (3.15)$$

then, by (3.11) and (3.15),

$$\alpha^{1/3} + \beta^{1/3} + \gamma^{1/3} = \theta = (a + 6 + 3t)^{1/3},$$

which proves (3.8). Also, by (3.13)–(3.15),

$$\varphi^3 = b - 3 + 3\theta\varphi = b + \theta^3 - a = b + 6 + 3t. \quad (3.16)$$

Hence, by (3.11) and (3.16), (3.9) is established. From (3.13) and (3.15),

$$3 + t = \theta\varphi. \quad (3.17)$$

Thus, by (3.15)–(3.17),

$$(3 + t)^3 = \theta^3\varphi^3 = (a + 6 + 3t)(b + 6 + 3t).$$

Expanding both sides, collecting terms, and simplifying, we deduce (3.10).

On page 356 of [26], the last page of the second notebook, Ramanujan offers the equalities

$$\left(\cos \frac{2\pi}{9}\right)^{1/3} + \left(\cos \frac{4\pi}{9}\right)^{1/3} - \left(\cos \frac{\pi}{9}\right)^{1/3} = \left\{\frac{3}{2}(9^{1/3} - 2)\right\}^{1/3} \quad (3.18)$$

and

$$\left(\sec \frac{2\pi}{9}\right)^{1/3} + \left(\sec \frac{4\pi}{9}\right)^{1/3} - \left(\sec \frac{\pi}{9}\right)^{1/3} = \{6(9^{1/3} - 1)\}^{1/3}, \quad (3.19)$$

which are applications of (3.8) and (3.9), respectively, with $a = 0$, $b = -3$, and $t = -9^{1/3}$. Equality (3.18) was posed as a problem by Ramanujan in the *Journal of the Indian Mathematical Society* [23], [27, p. 329]. Proofs of (3.18) and (3.19) can also be found in Berndt's book [5, Chapter 22].

4. NUMBER THEORY. Suppose p is a prime and n is a positive integer. Then, by a well-known theorem in elementary number theory [19, p. 182], the highest power of p dividing $n!$ equals

$$\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor =: N.$$

Despite the widespread use of this theorem by number theorists for many years, the inequalities

$$\frac{n}{p-1} - \frac{\log(n+1)}{\log p} \leq N \leq \frac{n-1}{p-1}, \quad (4.1)$$

given by Ramanujan [26, p. 378] in his third notebook do not appear to have been heretofore noticed. Both inequalities in (4.1) are sharp. If $n = p^m$ for some positive integer m , an elementary calculation shows that $N = (n-1)/(p-1)$.

On the other hand, if $n = p^{m+1} - 1$, by a direct calculation with the observation that $m + 1 = \log(n + 1)/\log p$,

$$N = \frac{n}{p-1} - \frac{\log(n+1)}{\log p}.$$

In fact, Ramanujan stated (4.1) with p replaced by an arbitrary positive integer $a \geq 2$.

Bhargava, Adiga, and Somashekara [7] have given one proof of (4.1) when p is any positive integer exceeding 1. We offer another proof here.

Proof of (4.1): First, by writing n in base p , i.e., by setting

$$n = \sum_{j=0}^m b_j p^j, \quad 0 \leq b_j \leq p-1, \quad b_m \neq 0,$$

we find, after a straightforward calculation, that

$$N = \frac{n}{p-1} - \frac{1}{p-1} \sum_{j=0}^m b_j, \quad (4.2)$$

and so the second inequality in (4.1) follows.

The first inequality in (4.1) is more difficult to establish. We are very grateful to B. Reznick for supplying the following elegant proof.

Set

$$b = \sum_{j=0}^m b_j.$$

Then, by (4.2), it suffices to prove that

$$b \leq (p-1) \frac{\log(n+1)}{\log p}. \quad (4.3)$$

Write

$$b = k(p-1) + r, \quad 0 \leq r \leq p-2. \quad (4.4)$$

Then

$$\begin{aligned} n &\geq (p-1)p^0 + (p-1)p + (p-1)p^2 + \cdots + (p-1)p^{k-1} + rp^k \\ &= (r+1)p^k - 1. \end{aligned}$$

It follows that

$$\begin{aligned} (p-1) \frac{\log(n+1)}{\log p} &\geq (p-1) \frac{\log((r+1)p^k)}{\log p} \\ &= k(p-1) + (p-1) \frac{\log(r+1)}{\log p}. \end{aligned} \quad (4.5)$$

By (4.3)–(4.5), we shall be finished with the proof if we can show that

$$r \leq (p-1) \frac{\log(r+1)}{\log p}. \quad (4.6)$$

First, if $r = 0$, (4.6) clearly holds with equality.

If $r \geq 1$, (4.6) can be written in the form

$$\frac{r}{\log(r+1)} \leq \frac{p-1}{\log p},$$

or

$$f(r) \leq f(p-1), \quad (4.7)$$

where

$$f(x) := \frac{x}{\log(x+1)}.$$

However, by elementary calculus, $f(x)$ is strictly increasing for positive integral x . Since $1 \leq r \leq p-2$, (4.7) is therefore valid with a strict inequality, and so the proof is complete.

As remarked in the Introduction, we conclude this short sampling of Ramanujan's elementary discoveries with a note on π . Continued fractions provide excellent rational approximations to π . Thus, the simple continued fraction

$$\pi = 3 + \frac{1}{7} + \frac{1}{15} + \frac{1}{1} + \frac{1}{293} + \cdots$$

yields the successive approximations $\frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \dots$. Note that

$$\frac{355}{113} = 3.14159\,29\dots,$$

which agrees with the decimal expansion of $\pi = 3.14159\,26535\dots$ through 6 decimal places. The appearance of a "large" fourth partial quotient, 293, is primarily responsible for this success.

Taking a brief diversion in his famous paper on approximations to π [24], [27, p. 35], Ramanujan offers the approximation

$$\pi \approx \left(97\frac{1}{2} - \frac{1}{11}\right)^{1/4} = 3.14159\,26526\dots, \quad (4.8)$$

which "was obtained empirically." How did Ramanujan deduce this unusual approximation, which is also found in his second and third notebooks [26, pp. 217, 375]? N. D. Mermin [16], [17, pp. 304–305] has offered the best explanation for Ramanujan's approximation (4.8). In the decimal expansion of $\pi^4 = 97.409091034002\dots$, observe that the pair of digits 09 appears twice in succession followed by the pair 10; which is 'close' to 09. Thus,

$$97.409090909\dots = \frac{2143}{22} = 97\frac{1}{2} - \frac{1}{11}$$

is a natural approximation to π^4 .

Ramanujan's facility with continued fractions is unequaled in mathematical history, and so he might have observed that [16], [17], [4, p. 151]

$$\pi^4 = 97 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{1} + \frac{1}{16539} + \frac{1}{1} + \cdots$$

Truncating this continued fraction just before the "super large" partial quotient 16,539 gives the approximation (4.8).

We are very grateful to Richard Askey, R. William Gosper, Daniel Grayson, K. Srinivasa Rao, and Bruce Reznick for valuable contributions.

REFERENCES

1. G. E. Andrews, R. A. Askey, B. C. Berndt, K. G. Ramanathan, and R. A. Rankin, editors, *Ramanujan Revisited*, Academic Press, Boston, 1988.
2. E. T. Bell, *The Last Problem*, Mathematical Association of America, Washington, D.C., 1990.
3. B. C. Berndt, *Ramanujan's Notebooks, Part II*, Springer-Verlag, New York, 1989.
4. ———, *Ramanujan's Notebooks, Part III*, Springer-Verlag, New York, 1991.
5. ———, *Ramanujan's Notebooks, Part IV*, Springer-Verlag, New York, 1994.
6. B. C. Berndt and S. Bhargava, A remarkable identity found in Ramanujan's third notebook, *Glasgow Math. J.* 34 (1992), 341–345.
7. S. Bhargava, C. Adiga, and D. D. Somashekara, An estimate of Ramanujan related to the greatest integer function, *Bull. Austral. Math. Soc.* 44 (1991), 149–154.
8. L. E. Dickson, *History of the Theory of Numbers*, vol. 2, Chelsea, New York, 1966.
9. N. D. Elkies, On $A^4 + B^4 + C^4 = D^4$, *Math. of Comp.* 51 (1988), 825–835.
10. L. Euler, *Solutio generalis quorundam problematum diophanteorum quae vulgo nonnise solutiones speciales admittere videntur*, Opera Omnia, Ser. I, vol. 2, B. G. Teubner, Lipsiae, 1915, pp. 428–458.
11. P. Fermat, *Oeuvres*, t. 3, Gauthier-Villars, Paris, 1896.
12. C. B. Haldeman, On biquadrate numbers, *The Mathematical Mag.* 2 (1904), 285–296.
13. G. H. Hardy, *Ramanujan*, 3rd edition, Chelsea, New York, 1978.
14. J. P. S. Kung and G.-C. Rota, The invariant theory of binary forms, *Bull. Amer. Math. Soc.* 10 (1984), 27–85.
15. A. Martin, About biquadrate numbers whose sum is a biquadrate-II, *The Mathematical Mag.* 2 (1904), 325–362.
16. N. D. Mermin, Pi in the sky, *Am. J. Phys.* 55 (1987), 584–585.
17. ———, *Boojums All the Way Through*, Cambridge University Press, Cambridge, 1990.
18. C. Moreau, Plus petit nombre égal à la somme de deux cubes de deux facons, *L'Intermédiaire Math.* 5 (1898), 66.
19. I. Niven, H. S. Zuckerman, and H. L. Montgomery, *An Introduction to the Theory of Numbers*, 5th ed., Wiley, New York, 1991.
20. S. Ramanujan, Question 284, *J. Indian Math. Soc.* 4 (1912), 183.
21. ———, Note on a set of simultaneous equations, *J. Indian Math. Soc.* 4 (1912), 94–96.
22. ———, Question 507, *J. Indian Math. Soc.* 6 (1914), 74–77.
23. ———, Question 524, *J. Indian Math. Soc.* 6 (1914), 190–191.
24. ———, Modular equations and approximations to π , *Quart. J. Math. (Oxford)* 45 (1914), 350–372.
25. ———, Question 722, *J. Indian Math. Soc.* 7 (1915), 240.
26. ———, *Notebooks* (2 volumes), Tata Institute of Fundamental Research, Bombay, 1957.
27. ———, *Collected Papers*, Chelsea, New York, 1962.
28. J. J. Sylvester, *Collected Mathematical Papers*, vol. 1, Chelsea, New York, 1973.
29. G. N. Watson, Solution to Question 722, *J. Indian Math. Soc.* 18 (1929), 113–117.
30. ———, Ramanujan's note books, *J. London Math. Soc.* 6 (1931), 137–153.

Department of Mathematics
University of Illinois
Urbana, IL 61801

Department of Mathematics
University of Mysore
Manasa Gangotri
Mysore 570 006, India

Chebychev Polynomials and Regular Polygons

D. Y. Savio and E. R. Suryanarayan

1. INTRODUCTION. Chebychev polynomials occur in many branches of mathematics: interpolation theory, orthogonal polynomials, approximation theory, numerical analysis, ergodic theory, etc. It is said that the Chebychev polynomial is like a fine jewel that reveals its different characteristics under illumination from varying positions [2]. There is yet a simple spot it shows its radiance: Regular Polygons. In this paper we study some of the properties of the Chebychev polynomials of the second kind $u_n(x)$, (2) below, and a polynomial associated with it, namely, $u_n + u_{n-1}$ and learn that the polynomials are related to some of the properties of the regular polygons. Specifically, we generalize a result due to Kepler (1571–1630). Kepler observed that the squares of the edges of polygons $\{7\}$, $\left\{\frac{7}{2}\right\}$, $\left\{\frac{7}{3}\right\}$ of unit circumradius (all having the same 7 vertices) are the roots of the equation,

$$z^3 - 7z^2 + 14z^2 - 7 = 0; \quad (1)$$

here, $\{7\}$ is a regular heptagon; $\left\{\frac{7}{2}\right\}$ and $\left\{\frac{7}{3}\right\}$ are star-polygons [1] (see FIGURE 1).

2. CHEBYCHEV POLYNOMIALS. Let $x = \cos \theta$. Chebychev polynomials of the second kind are defined recursively by [2], $u_n(x)$:

$$u_0(x) = 1, \quad u_1(x) = 2x, \dots, u_n(x) = \frac{\sin(n+1)\theta}{\sin \theta}, \quad n = 1, 2, \dots \quad (2)$$

$u_n(x)$ satisfies the classical recurrence

$$u_n(x) = 2xu_{n-1} - u_{n-2}. \quad (3)$$

The zeros of $u_n(x)$ are

$$\cos \frac{k\pi}{n+1}, \quad k = 1, \dots, n \quad (4)$$

[2]. We need the following result for later use.

Theorem 1. *Let $v_n(x) = u_n(x) + u_{n-1}(x)$. Then the zeros of $v_n(x)$ are $\cos(2k\pi/(2n+1))$.*

Proof: From the trigonometric definition of u_n and elementary trigonometric identities, we find that $v_n = (\sin(n + \frac{1}{2})\theta / \sin(\theta/2))$. Therefore, the zeros of $v_n(x)$ are $\cos(2k\pi/(2n+1))$. ■

Let

$$U_n(x) = \begin{bmatrix} 2x & 1 & 0 & 0 & \cdots & 0 \\ 1 & 2x & 1 & 0 & \cdots & 0 \\ 0 & 1 & 2x & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & 2x \end{bmatrix}, \quad (5)$$

where the matrix is symmetric, tridiagonal and is of order n . The result

$$u_n(x) = \text{Det } U_n(x) \quad (6)$$

follows from the recurrence (3) [3].

Theorem 2. The eigenvalues of $U_n(x)$ are $2(x - 1) + 4\sin^2(k\pi/2n + 2)$, $k = 1, 2, \dots, n$.

Proof: From (6) the eigenvalues of U_n are the zeros of $u_n(x - (\lambda/2))$; but from (4) the zeros of $u_n(x - (\lambda/2))$ are given $2x - 2\cos(k\pi/n + 1)$. Therefore, the eigenvalues of U_n are $2(x - 1) + 4\sin^2(k\pi/2n + 2)$. ■

There is a pair of recurrences less well known which also generates u_n :

$$u_{2n} = (u_n + u_{n+1})(u_{n+1} - u_{n-1}), \quad (7)$$

$$u_{2n+1} = u_n(u_{n+1} - u_{n-1}), \quad n = 1, 2, \dots \quad (8)$$

The preceding “odd-even” breakdown can be proved by using the definition (2) and elementary trigonometric identities. The above relations suggest that u_n is always the product of two determinants which are themselves “Chebychev polynomials” in the sense of $u_k \pm u_{k-1}$ and $u_{k+1} - u_{k-1}$ satisfy recurrence (3). Indeed,

$$\begin{aligned} u_k \pm u_{k-1} &= 2xu_{k-1} - u_{k-2} \pm u_{k-1} = (2x \pm 1)u_{k-1} - u_{k-2} \\ &= \text{Det} \begin{bmatrix} 2x & 1 & 0 & 0 & \cdots & 0 \\ 1 & 2x & 1 & 0 & \cdots & 0 \\ 0 & 1 & 2x & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & 2x \pm 1 \end{bmatrix} \end{aligned} \quad (9)$$

and

$$u_{k+1} - u_{k-1} = 2xu_k - 2u_{k-1} = \text{Det} \begin{bmatrix} 2x & 1 & 0 & 0 & \cdots & 0 \\ 1 & 2x & 1 & 0 & \cdots & 0 \\ 0 & 1 & 2x & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 2 & 2x \end{bmatrix} \quad (10)$$

differ from (5) just at a single entry.

When $n = 2k + 1$, let

$$V_n = \begin{bmatrix} 2x & 1 & 0 & 0 & \cdots & 0 \\ 1 & 2x & 1 & 0 & \cdots & 0 \\ 0 & 1 & 2x & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & 2x + 1 \end{bmatrix}, \quad n = 1, 3, 5, \dots \quad (11)$$

denote the square matrix of order k . When $n = 2k$, let U_n be defined by (5). All statements with U_n and V_n assume $n = 2k$ and $n = 2k + 1$ respectively.

Theorem 3. *The eigenvalues of $V_n(x)$ are $2(x - 1) + 4 \sin^2(k\pi/2n + 1)$, $k = 1, 2, \dots, n$.*

Proof: The eigenvalues of V_n are obtained by solving the equation $\text{Det } V_n(2(x - \lambda/2)) = 0$, for λ . Since $\text{Det } V_n = v_n$ and the zeros of $v_n(x - (\lambda/2))$ are $2(x - 1) + 4 \sin^2(k\pi/2n + 1)$, we see, from an argument similar to theorem 1, that the zeros of $v_n(x - (\lambda/2))$ are indeed the eigenvalues of V_n . ■

3. APPLICATION TO REGULAR STAR-FIGURES AND POLYGONS. A *regular star-figure* is a figure formed by connecting with straight lines every q th point, starting with one of the points that divide a circumference into n equal parts ($2q < n$); if all the n points are not connected, then start from the unconnected point next to the initial point, and repeat the connecting procedure until all the n points are connected. Such a star-figure is denoted by $\left\{ \frac{n}{q} \right\}$. If $q = 1$, we have a regular convex polygon $\{n\}$ of n sides. If n and q are relatively prime, the star-figure is a star-polygon or an n -gram. For a given n there are $\phi(n)/2$ regular n -grams where $\phi(n)$ is the Euler function, the number of numbers less than n and prime to it. If n and q are not relatively prime then $\left\{ \frac{n}{q} \right\}$ is a symmetrically superposed convex polygon. For example, $\left\{ \frac{5}{2} \right\}$ is a pentagram, whereas, the star-figure $\left\{ \frac{6}{2} \right\}$ (the star of David) is formed by two equilateral triangles symmetrically superposed.

Regular octagons and 16-gons occur in the mural decorations of ancient Egypt. Pentagrams and hexagons were used by the Babylonians. Pythagoreans used pentagram as a symbol of good health and also as a badge of recognition. Hindus use the star of Lakshmi $\left\{ \frac{8}{2} \right\}$ to symbolize the eight forms of wealth (*Ashtalakshmi*). Buddhists and Hindus draw, an elaborate form of star-figures and star-polygons, a *mandala*, on the ceremonial altars. The systematic study of the star-polygons was initiated and some of their properties were developed by Bradwardine (1290–1349), an English cleric who became Archbishop of Canterbury for the last month of his life.

Consider a regular star-figure of n sides. Let O be the center, M the mid-point and A one end of the side, and let $AM = (l/2)$ and $OA = R$ be the circumradius of the star-figure (FIGURE 2). The angle AOM is π/n for $\{n\}$ and $q\pi/n$ for the star-figure $\left\{ \frac{n}{q} \right\}$ and the edges are $2R \sin(\pi/2n)$ and $2R \sin(q\pi/2n)$ respectively.

From the above results and theorems 2 and 3, we have the following generalization of Kepler's observation:

Theorem 4. *Let l be an edge and R the circumradius of a regular star-figure of n sides. The eigenvalues of the matrices $V_n(1)$ and $U_n(1)$ are the ratios $(l/R)^2$ of the regular star-figures $\left\{ \frac{n}{j} \right\}$, $j = 1, 2, \dots, k$, if $n = 2k + 1$, and $j = 1, 2, \dots, k - 1$, if $n = 2k$.*

As a special case, consider the characteristic equation of $V_3(1)$:

$$\begin{bmatrix} 2 - \lambda & 1 & 0 \\ 1 & 2 - \lambda & 1 \\ 0 & 1 & 3 - \lambda \end{bmatrix} = 0.$$

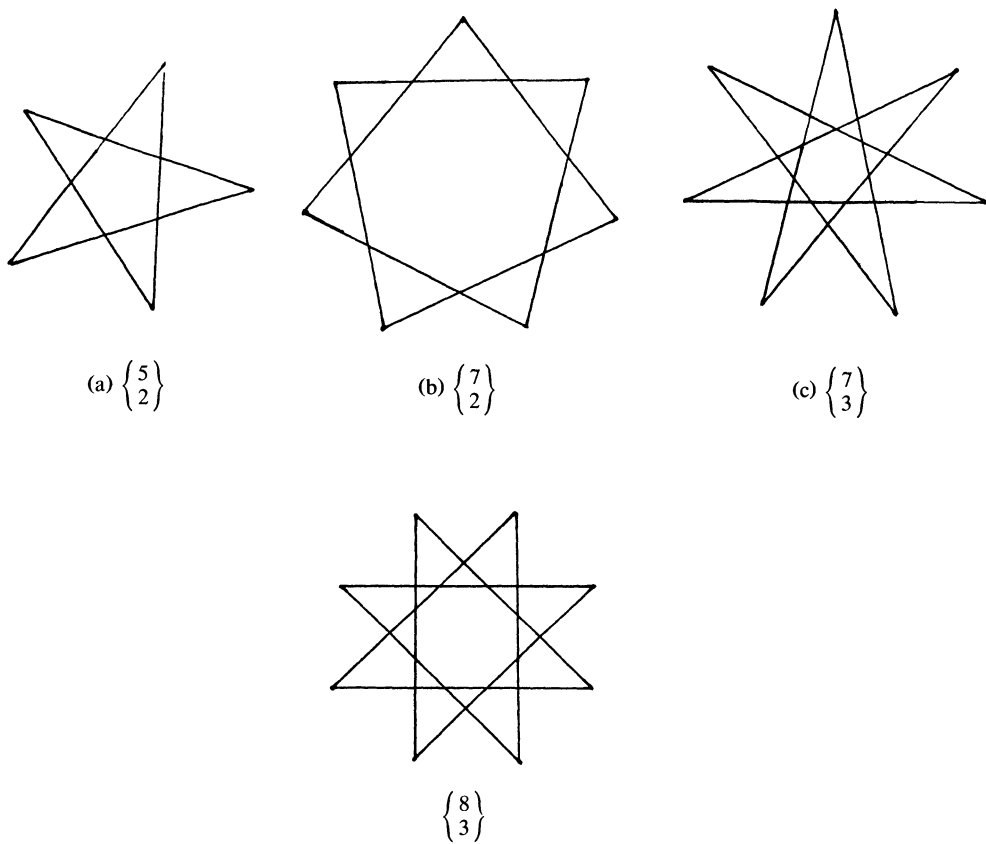


Figure 1

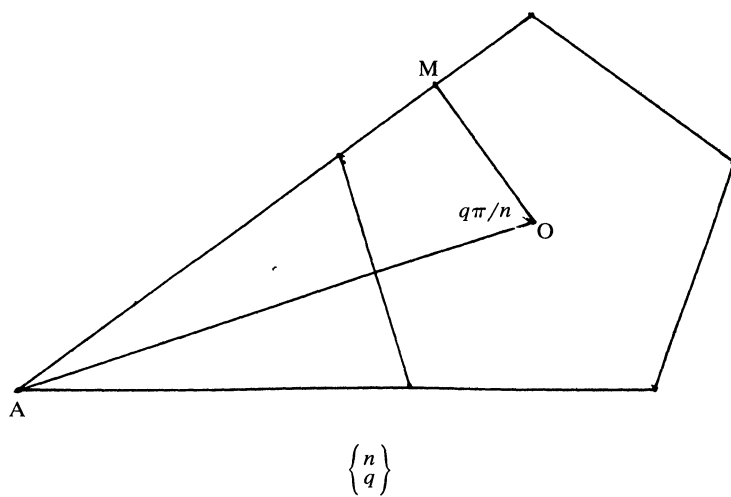


Figure 2

On expansion, the above equation reduces to $\lambda^3 - 7\lambda^2 + 14\lambda - 7 = 0$. Similarly, the three roots of the characteristic equation of $U_3(1)$:

$$\begin{bmatrix} 2 - \lambda & 1 & 0 \\ 1 & 2 - \lambda & 1 \\ 0 & 1 & 2 - \lambda \end{bmatrix} = 0,$$

give $(l/R)^2$ for $\{8\}$, $\begin{Bmatrix} 8 \\ 2 \end{Bmatrix}$, $\begin{Bmatrix} 8 \\ 3 \end{Bmatrix}$.

REFERENCES

1. H. S. M. Coxeter, *Regular Complex Polytopes*, Cambridge U. Press, 1974, p. 8.
2. T. J. Rivlin, *The Chebychev Polynomials*, John Wiley, New York.
3. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U. Press, 1985, p. 28.

Department of Mathematics
University of Rhode Island
Kingston, RI 02881

PICTURE PUZZLE

(from the collection of Paul Halmos)



A teacher's teacher---one of the greatest.
(see page 697.)

Small-Group Learning

Julian Weissglass

In any math class I've been in before, I just sat and listened to a teacher talk about what was in the book and what would be assigned for homework. When I got to college the same thing was happening except here I take notes on what the professor says. I have never felt in any of these situations that I should express my opinion on the subject. The most any teacher has done to stimulate a discussion on the topic was to simply say 'Questions?' Whenever the teacher said this, though, it didn't sound like he wanted a reply.

Junior Mathematics Major

In a span of less than two years, three national reports [6, 9, 10] have recommended fundamental changes in the teaching of college mathematics. The most recent document *Moving Beyond Myths* [10] states, for example, that "It is widely recognized that lectures place students in a passive role, failing to engage them in their own learning. Even students who survive such courses often absorb a very misleading impression of mathematics—as a collection of skills with no connection to critical reasoning" (p. 24). The document recommends that faculty, among other things, "explore effective alternatives to 'lecture and listen'", "involve students actively in the learning process," and "teach future teachers in the ways they will be expected to teach" (p. 34).

If we take seriously the charge to "teach future teachers in the ways they will be expected to teach," a reading of the *Professional Standards for Teaching Mathematics* [8] (which is referred to in *Moving Beyond Myths*) will lead to using small group approaches for at least part of the class time. This document states, "Students learning of mathematics is enhanced in a learning environment that is built as a community of people collaborating to make sense of mathematical ideas. It is a key function of the teacher to develop and nurture students abilities to learn with and from others—to clarify definitions and terms to one another, consider one another's ideas and solutions, and argue together about the validity of alternative approaches and answers . . ." (p. 58).

Reports and recommendations, of course, do not make changes in the classroom. Only teachers doing things differently achieve that. Changing teaching, however, is not easy. There is both *individual and institutional resistance* to change. My own experience with resistance occurred during my first attempt to use a small group approach in a linear algebra class I taught in 1970, my third year as a faculty member. Although the students liked the class I was so afraid that my colleagues would find out what I was doing that I closed the door of the classroom in case any of them walked by. My anxiety caused me to abandon the approach for three years.

At the institutional level, it was not until 1991 that the MAA annual meeting provided a special session devoted to alternatives to the lecture method, although articles [3, 11] appeared in the 70's describing this approach in mathematics courses and an increasing number of studies (see [4] for references) showed the effectiveness of small group cooperative learning approaches.

Having overcome, to some degree, my own resistance to pedagogical change, I thought it would be helpful to offer some suggestions to faculty considering

implementing small group approaches. In a sense, this is the article I wish I had been able to read twenty years ago.

BEGINNING. Do not be afraid to start slowly. It is not necessary to abandon lecturing completely. For some purposes it is a good method. You can combine lecturing, students working in groups, and whole-class discussion in any proportion you desire. One way to start is to have students form a group or pair and discuss how they solved a homework problem. Alternatively, you can pose an open-ended question for them to think about. Have them write about it (with a ‘quick write’) and then report their initial thinking to their group. Providing students time to think and write individually before sharing in the group is often helpful. Not all students want to start talking right away.

Another workable method is to set aside a portion of class time for students to discuss a concept or work on a problem, an investigation, or a group project. Some, or all, groups can report on their work to the class (either as a progress report or a final report). You can add perspective and background information as needed. In a large class, where it is cumbersome to use groups, you might have the students spend some time working in pairs—discussing a definition, sharing thoughts about a problem, comparing solutions or exploring a concept. Your Teaching Assistant can, with some encouragement from you, use small groups in discussion sections.

In order to ease the transition from lectures, provide an experience early on in the course demonstrating that a small group approach enhances learning in ways that lectures do not. For example, I often begin my class on problem solving with *Counting Squares*. Students are given a problem (FIGURE 1) and asked to work individually.

Counting Squares
(individual)

How many squares are there in the figure below? Be able to defend your answer. Work by yourself.

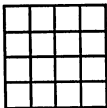


Figure 1

After about 10–15 minutes they are arranged in groups and given the problem in FIGURE 2.

Counting Squares
(small groups)

How many squares are there in the figure below? Work in your groups. Make sure that everyone is able to defend the answer.

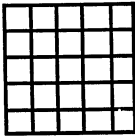


Figure 2

After the activity, I ask them to reflect on the process with the following instructions: Each person tells how they felt when doing the problem alone and as a group. Discuss the differences between individual learning and small group learning.

Another activity that shows students how small groups can enhance learning is Missing Corners. In this activity the students are asked to (individually) write a description of the pattern in FIGURE 3, construct with cubes (or tiles) the next two figures in the pattern, predict the number of cubes in the n th figure and write a justification for their prediction based on the figures. The students are amazed when they arrive at different ways of describing the pattern and justifying the predictions. (This can be followed by examining more complex, even 3-dimensional, patterns.)

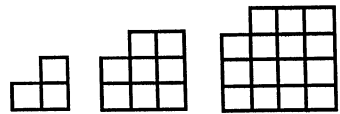


Figure 3

One obstacle when college students begin to work in groups is their lack of experience communicating about mathematics. It is important therefore that early group activities develop communication skills rather than stress solutions or proof. A colleague of mine, Bill Jacob, addresses this issue in a geometry class by having one student draw a geometric figure, a second write instructions on how to draw the figure for a third student, who then draws the figure without having seen the original figure. The figures are compared and the results discussed. Then the roles are rotated. He also has the students write reports on experiments (for example, projective geometry experiments with mirrors).

There are not many examples of college level curriculum written specifically for small group instruction. Two older texts [5, 12] attempt to present traditional course content for a small group approach. There is more available for the pre-college level and these sources may provide ideas for what can be done at the college level. The Interactive Mathematics Project¹ and the California Math A materials² are good examples of non-traditional approaches at the secondary level. Bishop [1] is a good resource for thinking about how to restructure curriculum for group projects and discussion. I used *Thinking Mathematically* [7] successfully in a problem solving course for potential secondary teachers. A good source for reading about what other people have done is [4]. Be aware, however, that some of the authors in this book have a very traditional view of mathematics and there is considerable disagreement about classroom practices as well.

It may be necessary to change your ideas of “covering” curriculum. It will help to reflect on the questions: what does it mean to teach? what does it mean to learn? College faculty need to think about and discuss the relative value of exposing students to mathematical knowledge or having them actually do mathe-

¹This project is developing a three year problem-based mathematics high school mathematics course. Contact Interactive Mathematics Project-EQUALS, University of California, Berkeley, CA 94720.

²This material was developed by California secondary teachers to meet the guidelines of the 1985 California Framework. It is being rewritten by Larry Hatfield for publication by Glencoe Publishing Company in June, 1993.

matics. For example, a class for potential secondary teachers explored symmetry by examining some strip patterns from Native American (San Ildefonso Pueblo) pottery. I then asked them to create their own strip patterns with pattern blocks (colored squares, triangles, trapezoids, parallelograms, and hexagons). I then asked the students to classify the strip patterns. With very little help from me most groups (in three to four hours of class time) discovered the seven different classes and were able to justify (although not rigorously) that these were all of them. I could have lectured about it in an hour or two, but I think that the level of understanding would have been shallower. There are no easy answers to questions about breadth versus depth—perhaps no answers at all—but it is beneficial to reflect on and discuss the questions.

STUDENTS. Students will probably be skeptical about participating in a small group at first. You need to explain to them why you are deviating from the traditional lecture method—and remind them periodically of your reasons and “philosophy of education”. Some students may continue to struggle with the different approach:

Once again as in Math 101A I have mixed feelings arising out of working in groups. One thing, working in small groups tends to make me more visible to others. That means my strong points, in between points, and weak points are right there for everyone to see. It is very hard for me to expose my weaknesses to others, i.e., my mistakes. Working in small groups tends to make me confront a feeling of stupidity. It is hard to overcome the urge to compare myself to others and to try to come up with the correct answer. I tend to underplay ‘correct’ contributions (i.e., a good idea) and overplay any errors I make, so it tends to be a struggle for me.

Others will make the transition more readily:

To be honest, when this class first began I did not enjoy it very much. It is hard for me to pin down why. In part it had a bit to do with the groups. It was not the fact that I did not know my group yet. I knew that we all would get to know one another. It was more because the people in my group seemed to be so much brighter than me. It did not seem as if I would ever have anything worth while to contribute After a few weeks of classes we all felt comfortable. We not only discussed math topics but also what we did over the weekend. How things were going etc. It was no longer a state of unfamiliarity or any anxiety over making a mistake or saying something foolish . . . [If we had not been in a group] I do not think we would have become friends.

Many come to understand and value the benefits of small group instruction:

Working in small groups is very different than lecturing only. There is no strict relationship where one person knows all the answers (teacher) and the other asks the questions (student). Working in a group is a more equal relationship where hopefully everyone is answering and asking questions. I like working in a small group, because it forces you to think rather than just copy whatever the teacher writes.

To be honest, in the past the only method that I knew to learn mathematics was to memorize so, therefore, if I memorized the material well I felt pretty good as a learner, but now, however, I realize that I have been somewhat cheated on what and how I learned mathematics. It just seems like I should have a better understanding of what I have learned in the past.

It is important to pay attention to the quality of the group process. Every three to four weeks I have students assess their group’s functioning. I ask them to answer two questions in their weekly journal: What are you doing to contribute to the group’s functioning well? What can you do to improve? Then I visit each group and sit down with them and ask each person to talk about their answers. This method provides them time to think about the questions free from pressure, but

ensures that the group is communicating about group processes. It also indicates clearly that I value group process, since I devote my time to assessing it.

Take time to interact with students. They will be uneasy about working in groups and will need time to talk about it. The relationship with the instructor is crucial in making the small group approach work. One student addressed the issue in his journal:

I believe that a very useful addition to this course would be to require, perhaps during the second or third weeks, each student to make an appointment during office hours. I believe a one on one discussion on ‘what do you want to get out of this course?’ to ‘what do you want to get out of teaching?’ would enhance the entire course. I believe that it would make the students even more aware of what they can get out of the course, as well as, being aware of the usefulness of being available to students for one-on-one talks. Offering offices hours does a lot. However, requiring us to take advantage of office hours would be excellent.

Be aware of the effect of grading on small groups. The first day (of a course in problem solving for prospective secondary teachers) I told the students how I would grade (see FIGURE 4) and in particular that I did not value memorization but would assess progress in their mathematical reasoning and their ability to communicate (verbally and in writing) about mathematics. I told them that I wanted to try (for the first time) using student portfolios to assess their work.

Attendance	10%;
Contribution to group and class (including an assessment of a portfolio of their work)	20%
Five problem sets	30%
Journal	20%
Final exam (oral)	20%

Figure 4

They were a little uneasy about this, so in the third week of the semester I spoke more about my philosophy, grades, and what portfolios were. I asked them to suggest what kind of evidence would show growth in mathematical thinking. We made a list and I indicated that they should include this type of evidence as part of their portfolio. Because I had devised what I thought would be a very acceptable grading method, and spent some time in class discussing it, I was surprised to read what one student wrote in her journal:

The assessment lecture bothered me. Up to that day working in the groups and learning was fun. I had been thoroughly enjoying the class but when the portfolio came up and I realized that something was going to be ‘graded’ my perspective of the class began to change. All of a sudden I had to pay attention to what I was writing down. ‘Is it neat enough?’ ‘Am I writing enough?’ ‘Have I misplaced something that I should have kept?’ These questions and slight panic began to be aroused in me. That day our group discussion was much more jumpy and less relaxed. For the first time our ideas came across in a competitive way. I cannot really explain why we became more interested in getting our ideas on paper than playing with the problem. With the knowledge that our progress was going to be measured, our performance became more forced and less enjoyable.

I have not solved the problem raised by this student. Certainly anxiety about grades is not unique to the small group approach. I have long believed that any “outside” (by someone other than the learner) evaluation of learning interferes with the learning process—with the possible exception of assessment conducted as an integral part of the learning process with the goal of assisting the learner. Furthermore I consistently find that my dual responsibilities of facilitating learning

and evaluating it, are inconsistent. Ideally a learner would be willing to reveal his/her ignorance to a teacher. In reality, he/she may be reluctant to do so to an evaluator. Although the small group approach reduces the interference of grades with learning (for example, anxiety is reduced by talking about grades with friends, students are graded on more than just test results) it does not eliminate it. I am not comfortable with grading and I admit my dilemma to my students. After reading the above student's journal, I read the passage (anonymously) to the class and we discussed grades, competition and learning. It seemed to help.

While on the subject of assessment, it is worth pointing out the obvious. Often students memorize to get by on tests. Success in this system does not necessarily mean that students have learned (understood and are able to use) the mathematics. We often do not notice this when using the lecture method because we only see test results. When you observe students working in groups, however, you will see more clearly what students do and do not understand. It can be disconcerting. In a class on classical number systems, for example, I asked students to use a concrete model to justify the familiar algorithm for adding fractions. They had tremendous difficulty coming up with an explanation.

A final point in regard to your students: Do not be too hard on them. There is an old saying "Don't blame the messenger who brings bad news." In a sense, undergraduates who cannot think or communicate well about mathematics are the message that something is drastically wrong with our education system. Small group instruction is not a panacea. It will not immediately remedy the deficiencies of previous miseducation. But it is a start.

INSTITUTIONAL AND PERSONAL SUPPORT. It will not be easy to give up the lecture method. Both institutional and personal support will be helpful in making the change. The Action Plan of *Moving Beyond Myths* makes many institutional recommendations. Draw these to the attention of relevant officials and organize on your campus for implementing the suggested reforms.

At present there is little opportunity for college faculty to participate in professional development focused on teaching. The educational community regards professional development in both content and methodology as a necessary part of *pre-college* teachers' professional growth. For college instructors, however, professional development focuses on learning more mathematics or on suggested revisions in content, not learning about new pedagogical approaches or research in mathematics education.

Until there is adequate opportunity to participate in professional development activities focused on teaching, individuals will have to strike out on their own. It may be possible to arrange your own professional development by watching someone who is using small groups or participating in a small group experience taught by someone else. In the long run, however, the attitudes and practices within the profession concerning professional development will need to change if large numbers of college faculty are to obtain the support necessary to implement the goals of the reform movement.

Even with institutional support for change you will need to get personal support if you intend to change your teaching. Find people with whom you can discuss mathematics teaching—your ideas, your successes and failures. (Accept that there will be failures.) In addition, find someone who is able to listen to you non-critically. It will be helpful to reflect on what you are doing and deal with your feelings about your efforts without fear of criticism. I did not have that 20 years ago and that is one reason why I abandoned my experimentation for three years. When you

are feeling tense or worried about whether you are doing the right thing you will tend to revert to the 'tried and true.' If you have someone to talk to about your feelings it is more likely that you will be able to think through the issues, and pursue your goals. See [13, 14] for further information about the relationship between feelings, listening and educational change.

CONCLUSION. Teaching using small groups is very different from lecturing. You will need time to develop your abilities. Be prepared for ambivalence and doubts. I encourage you to persist. Virtually every teacher (elementary or secondary) takes mathematics courses in a college or university. How you teach mathematics to undergraduates affects mathematics education throughout the entire system. You can play a crucial role in modeling for future teachers how to teach so that students are actively engaged in doing mathematics. If you are satisfied with large numbers of students not understanding or liking mathematics, with an attrition rate for mathematics students of approximately 50% each year after 9th grade [2], then continue with the lecture method. But if you want to provide opportunities for larger numbers of students to gain deeper understandings and to improve their ability to communicate about mathematics, then explore small group approaches and other alternatives to the lecture method. Perhaps you will be rewarded by having a future secondary teacher write: *I want to implement in my classroom what we did in this class. The most important thing that I have learned is that math can be fun.*

REFERENCES

1. Bishop, A. J., "Mathematical Enculturation: A Cultural Perspective on Mathematics Education." 1988 Kluwer. Dordrecht.
2. Committee on the Mathematical Sciences in the Year 2000. "A Challenge of Numbers." 1990 National Academy Press. Washington, D.C.
3. Davidson, N., The Small Group Discovery Method as Applied in Calculus Instruction. *American Mathematical Monthly*. August–September: 789–91, 1971.
4. Davidson, N., *Cooperative Learning in Mathematics, A Handbook for Teachers*. 1989, Addison Wesley, Menlo Park.
5. Davidson, N. and F. Gulick, "Abstract Algebra: An Abstract Learning Approach." 1976 Houghton Mifflin. Boston.
6. MAA Committee on the Mathematics Education of Teachers. "A Call for Change: Recommendations For The Mathematical Preparation Of Teachers Of Mathematics." Leitzel ed. 1991 Mathematical Association of America. Washington, D.C.
7. Mason, J., L. Burton and K. Stacey, "Thinking Mathematically." 1985 Addison Wesley. Reading.
8. National Council of Teachers of Mathematics. "Professional Standards for Teaching Mathematics." 1991 NCTM. Reston.
9. National Research Council. "Everybody Counts: A Report to the Nation About the Future of Mathematics Education." 1989 National Academy Press. Washington, D.C.
10. National Research Council. "Moving Beyond Myths: Revitalizing Undergraduate Mathematics." 1991 National Academy Press. Washington, D.C.
11. Weissglass, J., Small Groups: An Alternative to the Lecture Method. *Two Year College Mathematics Journal*. (Now the *College Mathematics Journal*) 7:15–20, 1976.
12. _____, "Exploring Elementary Mathematics: A Small Group Approach for Teaching," 1979 Kendall-Hunt (originally published by W. H. Freeman.)
13. _____, Constructivist Listening for Empowerment and Change. *The Educational Forum*. 54(4):351–370, 1990.
14. _____, Teachers Have Feelings: What Can We do About It? *Journal of Staff Development*. (12(1)):28–33, 1991.

*Mathematics Department
University of California
Santa Barbara, CA 93106*

A Fast Pick-Type Approximation for Areas of H -Polygons

Ding Ren, Krzysztof Kołodziejczyk, Grattan Murphy,
and John Reay

1. INTRODUCTION AND DEFINITIONS. Pick's formula $b/2 + i - 1$ gives the area of a simple polygon in R^2 whose corners lie in the integer lattice, and which has b lattice points on its boundary and i lattice points in its interior. It has been the object of many studies since its proof by Pick [6] in 1900. Throughout this paper we assume P is an H -polygon, i.e., a simple polygon whose corners lie in the set H of vertices of a monohedral tiling of R^2 by regular hexagons of unit area. See FIGURE 1 for examples. The vertices of this hexagonal tiling have density 2

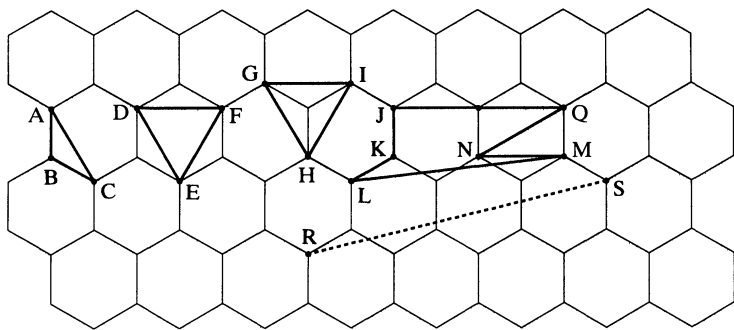


Figure 1. Measurable H -polygons.

(that is, each hexagon may be associated with 2 vertices), in contrast to the points of the integer lattice used in Pick's theorem, which have density one. Therefore it is reasonable to define *Pick's approximation* for the area $\mu(P)$ of an H -polygon P by

$$F(P) = (b/2 + i - 1)/2. \tag{1}$$

For example, in FIGURE 1 if P is the triangle ABC or triangle DEF then $b = 3$, and $i = 0$, so Pick's approximation is $F(P) = 1/4$, while the true areas are $\mu(ABC) = 1/6$ and $\mu(DEF) = 1/2$. Also triangle GHI has area $\mu(GHI) = 1/2$ and approximation $F(P) = (3/2 + 1 - 1)/2 = 3/4$. In the next section we find bounds on the size of the error of this Pick-type approximation for the area; this will show that $F(P)$ is, in some sense, a very good approximation.

The exact area of many H -polygons, like those in FIGURE 1, may be found by computing one additional parameter, the boundary characteristic. Every vertex $X \in H$ of the hexagonal tiling that is also on the boundary ∂P of P is the endpoint

of 3 edges of the hexagonal tiling. Define the *boundary characteristic* $c(X, P)$ of P at X to be the number of those 3 edges that extend locally into the exterior of P from X , minus the number that extend locally into the interior of P from X . Then the *boundary characteristic* of P is defined as $c = c(P) = \sum_{X \in H \cap \partial P} c(X, P)$. For example, if P is the irregular hexagon $JKLMNQ$ in FIGURE 1, then $b = 7$, $i = 0$, and $c = 2 - 1 + 2 + 3 - 3 + 3 + 1 = 7$. If points of H occur frequently along the boundary ∂P of P (specifically, if the neighboring H -points on ∂P are closer than the distance from R to S in FIGURE 1), then it is shown in [2] that the area $\mu(P)$ of P is given exactly by

$$A(P) = b/4 + i/2 + c/12 - 1. \quad (2)$$

(This is easily checked for the examples in FIGURE 1.) An H -polygon is called *measurable* if $\mu(P) = A(P)$. The measurable H -triangles have been characterized in [5]. Using the Pick approximation $F(P)$ for the area of P is faster than computing areas with (2) since it saves the computation of the boundary characteristic. In the next section we first get sharp inequalities between the parameters b and c for H -polygons, and then use them to justify the use of the fast Pick's approximation. Scott [7] and Coleman [1] have considered similar inequalities between b and i for convex polygons with corners in the integer lattice. See [4] and [8] for related results and further bibliography on Pick's Theorem.

2. BOUNDARY CHARACTERISTIC BOUNDS AND PICK'S APPROXIMATION.

Let P denote any simple H -polygon with $b = |H \cap \partial P|$ and boundary characteristic c . Triangles GHI and DEF of FIGURE 1, (and other examples with any $b \geq 3$) show that the inequalities in the following theorem are sharp.

Theorem 1. *For any simple H -polygon, $-b \leq c - 6 \leq b$.*

Theorem 1 may be used to provide a bound on the size of the error which occurs in using Pick's formula $F(P)$ to approximate the area $\mu(P)$ of measurable H -polygons.

Theorem 2. *If P is a measurable H -polygon then*

$$|F(P) - \mu(P)| \leq b/12.$$

Proof: If P is a measurable H -polygon then $A(P)$ in formula (2) gives the exact area $\mu(P)$ of P . Use the inequalities $c \leq b + 6$ and $c \geq -b + 6$ from Theorem 1 to replace c in the formula (2), and simplify. The result is immediate. ■

The triangles in FIGURE 1 show that the bound in Theorem 2 cannot be improved in general.

Proof of Theorem 1: We will choose a point in the relative interior of a side of P and traverse the boundary ∂P once in a counterclockwise direction, keeping track of changes in two parameters, the boundary characteristic and the deflection number, which will change only at points of $H \cap \partial P$. The *deflection number* as used in this proof will be defined in Table 1 for each $X \in H \cap \partial P$ in such a way

that it is always an integer multiple of $1/6$. The sum of the deflection numbers over all such X will agree with the *rotation number of P* as defined in [3] and [4], and will always be 1, which represents the fact that the direction of travel makes one complete rotation (of 2π) as we traverse once around ∂P in a counterclockwise direction. We may assume that the tilting at a typical vertex $X \in H \cap \partial P$ is oriented as shown in FIGURE 2, so that the 3 edges of the tiling which meet at X ,

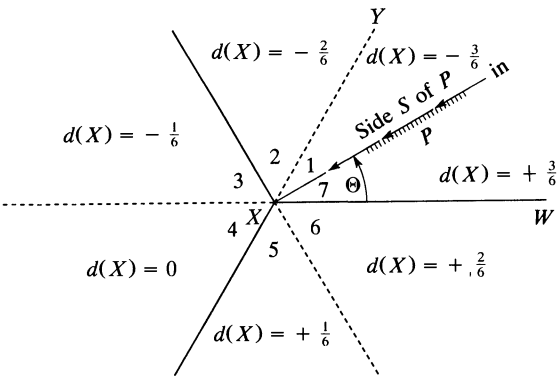


Figure 2. Sectors determined by a typical $X \in H \cap \partial P$.

together with their reflections in X , form 6 sectors about X , each of size $\pi/3$. Then the side S of P being traversed either approaches vertex X along the segment WX which is parallel to a tiling edge, or through the interior of sector WXY (as shown in FIGURE 2) thereby dividing sector WXY into 2 smaller sectors. In either case, let Θ be the angle between WX and side S , with $0 \leq \theta < \pi/3$. The sectors (numbered 1 through 7 in FIGURE 2) and the deflection number $d(X)$ for each sector are defined in Table 1.

TABLE 1. Deflection number when the boundary traverse leaves X in sector i . (Sector 7 does not exist if $\Theta = 0$ and side S contains WX .)

Sector Number	Angle ϕ from XW to leaving side	Deflection No. $d(X)$
1	$\Theta \leq \phi < \pi/3$	$-3/6$
2	$\pi/3 \leq \phi < 2\pi/3$	$-2/6$
3	$2\pi/3 \leq \phi < \pi$	$-1/6$
4	$\pi \leq \phi < 4\pi/3$	0
5	$4\pi/3 \leq \phi < 5\pi/3$	$1/6$
6	$5\pi/6 \leq \phi < 2\pi$	$2/6$
7	$0 \leq \phi < \Theta$	$3/6$

We distinguish three types of vertices $X \in H \cap \partial P$ depending on how the boundary passes through X on our traverse:

Type 1. The boundary traverse approaches X along side S with $\Theta > 0$ (as shown in FIGURE 2) and leaves X through the interior of some sector, or else, the traverse approaches X along WX (so $\Theta = 0$ and Sector 7 does not exist) and leaves X on a side parallel to an edge.

TABLE 2. Bounds for the boundary characteristics.

Sector Number	Type 1		Type 2		Type 3	
	c_M	c_m	c_M	c_m	c_M	c_m
1	-3	-3	(no such vertices)		-2	-3
2	-1	-3	-2	-3	-1	-2
3	-1	-1	-1	-2	0	-1
4	1	-1	0	-1	1	0
5	1	1	1	0	2	1
6	3	1	2	1	3	2
7	3	3	3	2	(no such vertices)	

Type 2. Angle $\Theta > 0$ and the traverse leaves X on a side which is parallel to an edge.

Type 3. The traverse approaches X along WX (so $\Theta = 0$) and leaves through the interior of some sector.

Table 2 shows the maximum $c_M(X, P)$ and minimum $c_m(X, P)$ possible values of $c(X, P)$, when vertex X is of each of the above 3 types. It will follow from the definition of the deflection number that

$$\sum_{X \in H \cap \partial P} d(X) = 1 \quad (3)$$

First, suppose that for each vertex X of P , our traverse of P always both approaches X and leaves X in the exact center of one of the 6 sectors of size $\pi/3$ shown in FIGURE 2. For this special case of P , the deflection number at each X agrees with the rotation number, takes a value from the discrete set $\{k/6 | k = -2, -1, \dots, +2\}$, and sums (over all $X \in H \cap \partial P$) to 1 full rotation. Hence (3) holds. To show (3) for a general H -polygon P , note that for each edge $\langle X, V \rangle$ of P the angle between $\langle X, V \rangle$ and the center of the sector it enters at X is exactly the negative of the angle between $\langle X, V \rangle$ and the center of the sector which it leaves at V . Thus the sum of the angle deflections is the sum of the deflection numbers and (3) holds for general P .

It is also clear that

$$c_L := \sum_{X \in H \cap \partial P} c_m(X, P) \leq c \leq \sum_{X \in H \cap \partial P} c_M(X, P) =: c_U \quad (4)$$

by the definition of the boundary characteristic. Define t_{ij} to be the cardinality of the set $\{X \in H \cap \partial P | \partial P \text{ leaves } X \text{ in sector } i, \text{ and } X \text{ is of type } j\}$ for $i = 1, 2, \dots, 7$. Then $b = \sum_{i \in \{1, 2, \dots, 7\}} \sum_{j \in \{1, 2, 3\}} t_{ij}$ and (3) may be rewritten (denoting $\sum_j t_{ij}$ by t_i) as

$$6 = -3t_1 - 2t_2 - t_3 + t_5 + 2t_6 + 3t_7 \quad (3')$$

and the right side of (4) becomes

$$\begin{aligned} c \leq & -3t_{11} - t_{21} - t_{31} + t_{41} + t_{51} + 3t_{61} + 3t_{71} \\ & - 2t_{22} - t_{32} + t_{52} + 2t_{62} + 3t_{72} \\ & - 2t_{13} - t_{23} + t_{43} + 2t_{53} + 3t_{63} = c_U. \end{aligned} \quad (4')$$

Using the above expressions for b , 6 , and c_U in terms of the t_{ij} 's, it follows that

$$6 + b = c_U + [t_{11} + t_{31} + t_{51} + t_{71}] \\ + t_{22} + t_{32} + t_{42} + t_{52} + t_{62} + t_{72} \geq c_U \geq c.$$

Using the left inequality of (4) in a similar way it follows that

$$6 - b = c_L - \left([t_{11} + t_{31} + t_{51} + t_{71}] + \sum_i t_{i3} \right) \leq c_L \leq c.$$

This gives the desired inequalities. ■

REFERENCES

1. D. B. Coleman, Stretch: a geoboard game, *Math. Mag.*, 51 (1978) 49–54.
2. R. Ding, K. Kołodziejczyk and J. Reay, A new pick-type theorem on the hexagonal lattice, *Discrete Math.* 68 (1988) 171–177.
3. B. Grünbaum and G. Shephard, Rotation and winding numbers for planar polygons and curves, *Trans. Amer. Math. Soc.* 322 (1990) 169–188.
4. B. Grünbaum and G. Shephard, Pick's theorem, *Amer. Math. Monthly* 100 (1993) 150–161.
5. K. Kołodziejczyk and J. Reay, Primitive and measurable hex-triangles, *Geometriae Dedicata* 15 (1992) 233–241.
6. G. Pick, Geometrisches zur Zahlenlehre, *Sitzungsber. Lotos. Prag.*, 19 (1900) 311–319.
7. P. R. Scott, On convex lattice polygons, *Bull. Austral. Math. Soc.*, 15 (1976) 395–399.
8. P. R. Scott, The fascination of the elementary, *Amer. Math. Monthly*, 94 (1987) 759–768.

Hebei Teachers University
Shijiazhuang, People's Republic of China

Instytut Matematyki
Politechniki Wrocławskiej
Wrocław, Poland

Mathematics Department
University of Maine
Orono, ME 04469

Mathematics Department
Western Washington University
Bellingham, WA 98225

Logarithmetica Britannica. Being a Standard Table of Logarithms to Twenty Decimal Places. By Alexander John Thompson. Part V, Numbers 50000 to 60000 Issued by the Biometric Laboratory, University of London, to Commemorate the Tercentenary of Henry Briggs' Publication of the *Arithmetica Logarithmica*, 1624. Subscription Issue. Cambridge, The University Press, 1931.

This is the fifth part (the fourth not yet published) of this tremendous undertaking. It consists of twenty-place logarithms of numbers of five digits, accompanied by values of second and fourth differences. The project speaks for itself; it is sufficient to say that the result is all that is to be expected of any product of the Cambridge Press.

—*American Mathematical Monthly*
38, (1931) p. 407

NOTES

Edited by: John Duncan

A Simple Example on Non-Sequentialness in Topological Spaces

Heinz König

There are well-known examples of countable Hausdorff topological spaces which are not discrete but show certain typical features of discreteness: all compact subsets are finite, and therefore all convergent sequences are ultimately constant. These spaces are of interest in functional analysis and measure theory. The examples known to the present author are due to Arens [1] and Varadarajan [6]. This note wants to add a different example which is strikingly simple. It arose in [5] in connection with the double limit relation.

THE NEW EXAMPLE. We fix a sequence (t_n) of real numbers $t_n > 0$ such that $t_n \rightarrow 0$ for $n \rightarrow \infty$ and $\sum_{n=1}^{\infty} t_n = \infty$ (for example $t_n = 1/n$). We call a subset $S \subset \mathbb{N}$ *small* iff $\sum_{n \in S} t_n < \infty$, which is to include $S = \emptyset$. Thus all finite subsets of \mathbb{N} are small, but there are also infinite small subsets: in fact, each infinite subset of \mathbb{N} contains a small infinite subset. By means of the small subsets of \mathbb{N} one then forms a topology on $X := \mathbb{N} \cup \{\infty\}$, with the open sets defined to be i) all subsets $A \subset \mathbb{N}$, and ii) those subsets $A \subset X$ with $\infty \in A$ whose complements $A' \subset \mathbb{N}$ are small. It is obvious that this is a Hausdorff topology on X which is not discrete, and it is a simple verification that all compact subsets of X are finite.

We turn to the two previous examples. Each time the above rôle of the small subsets of \mathbb{N} will be assumed by some other set system σ on \mathbb{N} .

The example of Arens (see also Kelley [4] Problem 2.E and Engelking [3] Example 1.6.20). We fix a sequence (X_n) of pairwise disjoint infinite subsets $X_n \subset \mathbb{N}$ with union \mathbb{N} . Then we define σ to consist of the subsets $S \subset \mathbb{N}$ such that $S \cap X_n$ is finite for almost all n .

The example of Varadarajan (see also Berg-Christensen-Ressel [2] Exercise 2.1.30). We define σ to consist of the subsets $S \subset \mathbb{N}$ such that

$$\frac{1}{n} \text{card}(S \cap \{1, \dots, n\}) \rightarrow 0 \quad \text{for } n \rightarrow \infty.$$

Thus each time we have a system σ of subsets of \mathbb{N} , intended to form the *small* subsets of \mathbb{N} , with the properties

- 1) σ contains all finite subsets of \mathbb{N} ;
- 2) $S \in \sigma$ implies $T \in \sigma$ for all $T \subset S$;

- 3) σ is stable under finite unions;
- 4) \mathbb{N} is not a member of σ ;
- 5) each infinite $S \subset \mathbb{N}$ contains an infinite $T \in \sigma$.

In each of the two previous examples properties 1)–4) are obvious and 5) requires a little proof. In the new example all properties are obvious.

By means of a set system σ on \mathbb{N} with the above properties 1)–5) one then forms a topology τ on $X := \mathbb{N} \cup \{\infty\}$, with the open sets defined to be i) all subsets $A \subset \mathbb{N}$, and ii) those subsets $A \subset X$ with $\infty \in A$ whose complements $A' \subset \mathbb{N}$ are members of σ . We collect the main consequences in the proposition below, the proof of which can be left to the reader as a sequence of simple exercises. We note that the last assertion is independent of condition 5) above.

Proposition. 1) τ is a Hausdorff topology on X which is not discrete. 2) Each compact subset (and even each relatively countably compact subset) of X is finite. Therefore each convergent sequence in X is ultimately constant. 3) τ is completely normal: each nonvoid subset of X is normal in its relative topology.

In particular the subset $\mathbb{N} \subset X$ has the cluster point ∞ . Also the sequence (x_n) of the points $x_n = n$ has the cluster value ∞ . But there is no sequence in \mathbb{N} which converges to ∞ .

Thus we have a common scheme for all the above examples. It is obvious that the new example is particularly simple.

There is also the notorious non-constructive example: By Zorn's lemma, each set system σ on \mathbb{N} with properties 1)–4) (for example the system of all finite subsets) is contained in a maximal such set system (in order to work with the usual notions of filters and ultrafilters one has to pass to complements). We claim that each set system σ on \mathbb{N} which is maximal with respect to properties 1)–4) also satisfies 5), and hence produces a topology τ on $X := \mathbb{N} \cup \{\infty\}$ as above. To see this note first that for each $T \subset \mathbb{N}$ one has either $T \in \sigma$ or $T' \in \sigma$. Now fix an infinite $S \subset \mathbb{N}$. We write $S = P \cup Q$ with disjoint infinite $P, Q \subset S$. In case $P, Q \notin \sigma$ then $P', Q' \in \sigma$ and hence $\mathbb{N} = (P \cap Q)' = P' \cup Q' \in \sigma$, which is not true. Thus we have a $P \in \sigma$ or $Q \in \sigma$. This proves 5).

The author wants to thank the referee and the Notes editor for good advice.

REFERENCES

1. R. Arens, Note on convergence in topology. *Math. Mag.* 23 (1950), 229–234.
2. J. Berg, J. P. R. Christensen and P. Ressel, *Harmonic Analysis on Semigroups*, Springer, 1984.
3. R. Engelking, *General Topology*, Heldermann, Berlin, 1989.
4. J. L. Kelley, *General Topology*, Van Nostrand 1955.
5. H. König and N. Kuhn, Angelic spaces and the double limit relation, *J. London Math. Soc.* (2) 35 (1987), 454–470.
6. V. S. Varadarajan, Measures on topological spaces, *AMS Transl.* (II) 48 (1965), 161–228.

Universität des Saarlandes
Fachbereich Mathematik
D-66041 Saarbrücken / Germany

The Secant Method and the Golden Mean

Melvin J. Maron and Robert J. Lopez

The secant method is a well-known method for finding roots α of the equation $f(x) = 0$. Starting with two initial approximations of α , say x_{-1} and x_0 , the secant method generates

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}, \quad k = 0, 1, 2, \dots \quad (1)$$

The rate at which the sequence $\{x_k\}$ converges to a root α depends on the multiplicity of α . Recall that α is a root of *multiplicity* m of the function f if $f(x)$ can be written as

$$f(x) = (x - \alpha)^m \phi(x), \quad \text{where } \phi \text{ is bounded at } \alpha \text{ and } \phi(\alpha) \neq 0. \quad (2)$$

It is well known (see [2]) that if α is simple root, that is, if $m = 1$, then there will be a nonzero asymptotic error constant C such that the errors

$$\varepsilon_k = \alpha - x_k$$

satisfy

$$\lim_{k \rightarrow \infty} \frac{|\varepsilon_k|}{|\varepsilon_{k-1}|^p} = C, \quad \text{where } p = \frac{\sqrt{5} + 1}{2} = 1.618 \dots$$

Thus, secant method iterates will converge superlinearly with order $p = \frac{1}{2}(\sqrt{5} + 1)$ to simple roots α . Ancient Greek mathematicians attached profound significance to the numbers

$$r = \frac{\sqrt{5} - 1}{2} = 0.618 \dots \quad \text{and} \quad p = r + 1 = \frac{1}{r} = \frac{\sqrt{5} + 1}{2} = 1.618 \dots \quad (3)$$

They referred to r as the *golden mean* because the ratio $(1 - r):r$ equals r . Observe that r and $-p$ are the roots of the quadratic equation

$$x^2 + x - 1 = 0. \quad (4)$$

The purpose of this paper is to prove the following result¹, which shows that the golden mean is also related to the way secant method approximants converge to double roots.

Theorem. Suppose α is a root of f for which

$$f(x) = (x - \alpha)^2 \phi(x), \quad \text{where } \lim_{x \rightarrow \alpha} \phi(x) \neq 0, \quad (5)$$

¹Part (a) was obtained for $f(x) = x^2(x - 1)^2$ and $\alpha = 1$ in [1] (Example E-11, p. 278).

and let $\varepsilon_k = \alpha - x_k$, where x_k is the k th approximant generated by secant method (1).

(a) If the sequence $\{x_k\}$ converges to α and $\lim_{k \rightarrow \infty} (\varepsilon_k / \varepsilon_{k-1})$ exists, then

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_k}{\varepsilon_{k-1}} = r, \quad \text{where } r = \frac{\sqrt{5} - 1}{2} = 0.618 \dots$$

(b) For x_{k-1} , $x_k \approx \alpha$ and $\rho_k = \varepsilon_k / \varepsilon_{k-1} \approx r$, the ratios ρ_k will satisfy

$$\rho_{k+1} - r \sim -r^2(\rho_k - r) + r \left[\frac{\phi(x_{k-1})}{\phi(x_k)} - 1 \right] \quad (6)$$

It follows that $\{\rho_k\}$ will converge to r once x_{k-1} and x_k are sufficiently close to α and ρ_k is sufficiently close to r .

Proof: (a) Since $x_k - x_{k-1} = (x_k - \alpha) + (\alpha - x_{k-1}) = \varepsilon_{k-1} - \varepsilon_k$, we have from (1)

$$\varepsilon_{k+1} = \alpha - x_{k+1} = \varepsilon_k + \frac{f(x_k)(\varepsilon_{k-1} - \varepsilon_k)}{f(x_k) - f(x_{k-1})}. \quad (7)$$

In view of (5), we may assume x_{k-1} and x_k to be sufficiently close to α so that ε_k , $\phi(x_k)$, and $f(x_k) = \varepsilon_k^2 \phi(x_k)$ are all nonzero. Under this assumption, we can rearrange (7) to get

$$\frac{\varepsilon_{k+1}/\varepsilon_k - 1}{1 - \varepsilon_{k-1}/\varepsilon_k} = \frac{1}{f(x_{k-1})/f(x_k) - 1}, \quad (8)$$

where $f(x_{k-1})/f(x_k) = \phi(x_{k-1})/\phi(x_k) \cdot (\varepsilon_{k-1}/\varepsilon_k)^2$. Upon introducing the ratios

$$\rho_k = \frac{\varepsilon_k}{\varepsilon_{k-1}} \quad \text{and} \quad \beta_k = \frac{\phi(x_{k-1})}{\phi(x_k)}, \quad \text{for } k = 1, 2, \dots$$

in (8) and using the assumption $\lim_{x \rightarrow \alpha} \phi(x) \neq 0$, we get

$$\frac{\rho_{k+1} - 1}{\rho_k - 1} = \frac{\rho_k}{\beta_k - \rho_k^2}, \quad \text{where } \lim_{k \rightarrow \infty} \beta_k = 1. \quad (9)$$

So if $\lim_{k \rightarrow \infty} \rho_k$ exists, it must be a solution of the equation $(x - 1)(x^2 + x - 1) = 0$, that is 1, r , or $-p$. Since $\lim_{k \rightarrow \infty} \rho_k$ cannot be zero, $\{x_k\}$ cannot converge superlinearly to α ; however, linear convergence requires $|\lim_{k \rightarrow \infty} \rho_k| \leq 1$. But $\lim_{k \rightarrow \infty} \rho_k$ cannot be 1 because if it were, then

$$|\rho_{k+1} - 1| < |\rho_k - 1|, \quad \text{that is, } \left| \frac{\rho_{k+1} - 1}{\rho_k - 1} \right| < 1$$

would hold for infinitely many k 's in (9), whereas $|\rho_k/(\beta_k - \rho_k^2)| > 1$ would hold for sufficiently large k . This leaves r as the only possible asymptotic error constant.

(b) To obtain (6), we first rewrite (9) as the finite difference equation

$$\rho_{k+1} = \frac{\rho_k - \beta_k}{\rho_k^2 - \beta_k}, \quad k = 0, 1, 2, \dots \quad (10)$$

and then subtract r to get

$$\begin{aligned}
 \rho_{k+1} - r &= \frac{\rho_k - \beta_k}{\rho_k^2 - \beta_k} - r \\
 &= \frac{\rho_k - \rho_k^2 r - \beta_k(1 - r)}{\rho_k^2 - \beta_k} \\
 &= \frac{\rho_k^2 r - \rho_k + \beta_k r^2}{\beta_k - \rho_k^2}. \quad [1 - r = r^2]
 \end{aligned}$$

Writing ρ_k as $(\rho_k - r) + r$ and collecting numerator terms gives

$$\begin{aligned}
 \rho_{k+1} - r &= \frac{(\rho_k - r)^2 r + (\rho_k - r)(2r^2 - 1) + (r^3 - r + \beta_k r^2)}{\beta_k - \rho_k^2} \\
 &= \frac{r}{\beta_k - \rho_k^2} \{(\rho_k - r)(\rho_k - 1) + r(\beta_k - 1)\} \quad [r^2 - 1 = -r] \\
 &= \frac{r(\rho_k - 1)}{\beta_k - \rho_k^2} (\rho_k - r) + \frac{r^2}{\beta_k - \rho_k^2} (\beta_k - 1). \quad (11)
 \end{aligned}$$

Observe that if $\beta \rightarrow 1$ and $\rho \rightarrow r$, then

$$\frac{r(\rho - 1)}{\beta - \rho^2} \rightarrow \frac{r(r - 1)}{1 - r^2} = \frac{-r}{1 + r} = -r^2$$

[see (3)] and, similarly, $r^2/(\beta - \rho^2) \rightarrow r^2/(1 - r^2) = r$. It thus follows from (11) that

$$\rho_{k+1} - r = (-r^2 + \mu_1)(\rho_k - r) + (r + \mu_2)(\beta_k - 1)$$

where $|\mu_1|$ and $|\mu_2|$ can be made arbitrarily small by keeping $|x_k - \alpha|$ and $|\rho_k - r|$ sufficiently small. This implies (6) and completes the proof of the theorem. \square

Traub [1, p. 278] states that secant method iterates will converge linearly to roots of any multiplicity $m > 1$. However, the authors are aware of no proof of this plausible assertion in [1] or elsewhere.

REFERENCES

1. J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice Hall, Englewood Cliffs New Jersey, 1964.
2. M. Vianello and R. Zanollo, On the Superlinear Convergence of the Secant Method, *American Mathematical Monthly*, 99 8 (1992) 758–761.

*Engineering Mathematics and
Computer Science Department
University of Louisville
Louisville, KY 40292
mjmaro01@ulkyvx.louisville.edu*

*Mathematics Department
5500 Wabash Avenue
Rose-Hulman Institute of Technology
Terre Haute, IN 47803
lopez@rose-hulman.edu*

R_n Contains a Division Ring iff R Does

Ayman Badawi

INTRODUCTION. Let R be a ring with 1, and let R_n denote the complete matrix ring of all $n \times n$ matrices over R under the usual matrix addition and multiplication. Recall $A, B \in R_n$ are similar iff there exists $P \in R_n$ such that $A = PBP^{-1}$. If $A \in R_n$ is similar over R to a diagonal matrix, then A is called [1] diagonalizable over R . For $B \in R_n$, b_{ij} denotes the entry of B in the i th row and j th column.

In this note, we give an alternative proof of [1, Theorem 1] which is quite shorter than that in [1]. We would like to point out that our proof begins exactly like the original.

Theorem ([1, Theorem 1]). *Let R be a ring with 1 for which each idempotent matrix in R_n is diagonalizable over R . Then R contains a division ring if and only if R_n contains a division ring.*

Proof: If R contains a division ring, then clearly R_n contains a division ring. Assume R_n contains a division ring K . The division ring K has an identity—call it J —and by the hypothesis $PJP^{-1} = I$ a diagonal matrix for some invertible matrix $P \in R_n$. Since the conjugation of R_n by P induces a ring automorphism of R_n , $M = PKP^{-1}$ is a division ring of R_n and has I as the identity. Hence I is a nonzero idempotent of R_n . Let $S = \{A \in M: A \text{ is diagonal}\}$. Since $I \in S$, S is not empty. We leave it to the reader to verify that S is a division subring of M . Since $I \neq 0$, there exists $1 \leq j \leq n$ such that i_{jj} is a nonzero idempotent of R . Let $D = \{a_{jj}: A \in S\}$. Then D is a division ring of R with i_{jj} as the identity.

We end this note with some examples that satisfy the hypothesis of the Theorem and with one example where the hypothesis fails. Let R be a commutative ring with 1. Then R is called *ID* (basal) as in [7] ([2]) iff for every $n \geq 1$ the idempotents of R_n are diagonalizable. Foster [2] has shown that if R is a principal ideal domain, then R is *ID*. Seshadri [6] has shown that if R is a principal ideal domain, then $R[x]$ is *ID*. In particular if F is a field, then $F[x, y]$ is *ID*. Steger [7] has shown that if R is an elementary division ring (i.e., for every $n \geq 1$ and $A \in R_n$ there exist invertible matrices P, Q in R_n such that PAQ is diagonal) then R is *ID*. Also; Steger has shown that if R is π -regular ring (i.e., for every x in R there exists $n \geq 1$ and y in R (n and y depending on x) such that $x^n y x^n = x^n$) then R is *ID*. In particular for every $m \geq 1$ Z_m (i.e., Z/mZ) is *ID* (Foster has shown independently that Z_m is *ID*).

Finally, Theorem 3 in [7] states that if R is *ID*, then every invertible ideal of R is principal. Thus if R is a Dedekind domain which is not principal, then R is not *ID*. In particular, let $R = Z[\sqrt{-5}]$ (Z is the set of all integers). Then R is a Dedekind domain, see [4, EX. 37, P. 70]. But R is not a unique factorization domain, for example 21 does not have unique factorization in R . Thus R is not principal and therefore it is not *ID*.

REFERENCES

1. Jacob T. B. Beard, Jr. and Robert McConnel, Matrix fields over the integers modulo m , *Linear Algebra and Its Applications* 14, (1976), 95–105.
2. A. L. Foster, Maximal idempotent sets in a ring with unit, *Duke Math. J.* 13 (1946) 247–258.
3. L. Gilman and M. Henriksen, Some remarks about elementary divisor rings, *Trans. Amer. Math. Soc.* 82 (1956), 362–365.
4. Harry C. Hutchins, *Examples of Commutative Rings*, Harry C. Hutchings, 1981.
5. I. Kaplansky, Elementary divisors and modules, *Trans. Amer. Math. Soc.* 66 (1949), 464–491.
6. C. S. Seshadri, Triviality of vector bundles over affine space K_2 , *Proc. Nat. Acad. of Sci. USA* 44 (1958), 456–458.
7. A Steger, Diagonability of idempotent matrices, *Pac. J. Math.* 19 (1966), 535–542.

Dedicated to Prof. Nick Vaughan on his retirement.

Department of Mathematics
University of North Texas
Denton, TX. 76203

A Further Simplification of Dixon's Proof of Cauchy's Integral Theorem

Peter A. Loeb

The modification in [1] of Dixon's proof of the Cauchy Integral Theorem and Formula is based on the proposition stated below. In this note we give a proof of that proposition which is more suitable for undergraduate students. In what follows, G will be an open set in the complex plane \mathbb{C} , and γ will be a closed rectifiable curve. We write $f \in H(G)$ if f is holomorphic, i.e. analytic, in G , and we use the notation $D(z, r)$ for the disk $\{w \in \mathbb{C}: |w - z| < r\}$. The trace of γ in \mathbb{C} is denoted by $\{\gamma\}$; we say the curve γ is in G when $\{\gamma\} \subset G$.

Proposition. *If γ is a curve in G , then for any $z \in \{\gamma\}$ there is a closed curve σ in G with $z \notin \{\sigma\}$ such that $\int_\gamma f = \int_\sigma f$ for all $f \in H(G)$.*

Proof: We assume that there is a point $\zeta \neq z$ with $\zeta \in \{\gamma\}$; otherwise the result is trivial. Pick $r > 0$ so that $D(z, r) \subset G$ and $\zeta \notin D(z, r)$. We will assume that γ is given by $\gamma(t)$ for $t \in [0, 1]$ and $\gamma(0) = \gamma(1) = \zeta$. By the uniform continuity of the mapping γ , there is a natural number n such that if $s, t \in [0, 1]$ and $|t - s| < 1/n$, then $|\gamma(t) - \gamma(s)| < r$. Partition the interval $[0, 1]$ using the points $0 < 1/n < \dots < (n-1)/n < 1$. Let $0 = x_0 < x_1 < x_2 < \dots < x_m = 1$ be the set of partition points k/n such that $\gamma(k/n) \neq z$. If between adjacent points x_i and x_{i+1} there is a point of the form k/n or any other point t_0 with $\gamma(t_0) = z$, then the path $\gamma(t)$, $x_i \leq t \leq x_{i+1}$, is in the disk $D(z, r)$. In this case, we may replace the

map γ on the interval $[x_i, x_{i+1}]$ with a path that goes from $\gamma(x_i)$ to $\gamma(x_{i+1})$ in the set $D(z, r) - \{z\}$. By Cauchy's integral theorem, applied to the disk $D(z, r)$, this replacement does not change the value of the integral for any $f \in H(G)$. With these replacements, the new path σ avoids z . \square

REFERENCE

1. P. A. Loeb, A note on Dixon's proof of Cauchy's Integral Theorem, *American Mathematical Monthly* 98 (1991) 242–244.

Department of Mathematics
University of Illinois
1409 West Green St.
Urbana, Ill. 61801
LOEB@MATH.UIUC.EDU

Who Was the Author?

On an extension of Sir John Wilson's theorem to all numbers whatever,
Philosophical Magazine, 1839.

Note on elimination, *Philosophical Magazine*, 1840.

Proof of the hitherto undemonstrated fundamental theorem of invariants,
Philosophical Magazine, 1878.

On a point in the theory of vulgar fraction, *Amer. Jour. of Math*, 1880.

On Farey Series, Johns Hopkins University Circulars, 1883.

Answer on page 697.

Zero-Knowledge Proofs

Catherine C. McGeoch

On a moonless night the spy returns to the castle after a reconnoitering mission to the enemy camp. As he nears the gate a voice whispers, “What’s the password?” But is it friend or foe who whispers? How can the spy show that he knows the password without actually revealing it to a possible imposter?

The spy’s dilemma is commonplace now with the widespread use of telecommunications. When your automatic teller machine communicates with your bank, each must be assured that the other is legitimate; the electronic “passwords” must be unforgeable and must be of no use to imposters and eavesdroppers. One method that has been proposed for exchanging passwords in this context is the *zero-knowledge proof*.

Renaissance mathematicians developed their own primitive zero-knowledge proof systems. When both Tartaglia and Fior claimed knowledge of an algebraic solution to cubic equations, a contest was arranged in which each proposed thirty problems for the other to solve. In the end, Tartaglia had solved all thirty, thus providing a convincing demonstration that he knew the method without actually revealing it. Fior solved none. (It turned out that each had worked out solutions to certain classes of cubics, but neither had solved the general problem [2]).

We’ve progressed considerably in formalizing this idea. An *interactive protocol* comprises two algorithms P (the prover) and V (the verifier) that read a common *input string* w of length $|w|$ and then compute and communicate in alternating turns to determine whether w has some specified property. The verifier is *polynomially-bounded*: it must eventually halt and its total computation time must be bounded by a fixed polynomial in $|w|$. When it halts, the verifier outputs either *accept* or *reject* depending upon whether the property holds for w . The verifier is *probabilistic*, that is, allowed to make random choices during the computation according to the results of coin tosses. The prover is allowed to have unlimited computational power.

A *language* \mathcal{L} is a set of strings. An *interactive proof system* for \mathcal{L} is an interactive protocol in which P helps V to decide whether $w \in \mathcal{L}$. We require that with high probability the verifier be correct when accepting or rejecting the membership of w in \mathcal{L} . More precisely, for every constant $c > 0$, for sufficiently large $w \in \mathcal{L}$ the probability (over all coin tosses) that V halts and accepts must be at least $1 - |w|^{-c}$. If $w \notin \mathcal{L}$ then we require that no prover P^* be able to convince V otherwise: that is, for every $c > 0$ and large enough w , and for any interactive protocol (P^*, V) , V rejects with probability at least $1 - |w|^{-c}$.

In a *zero-knowledge* interactive proof system, whenever $w \in \mathcal{L}$, P reveals no additional knowledge beyond the fact of membership. Informally, “no additional knowledge” means that the computational power of any verifier V^* after participating in the protocol is no more than what V^* would have gained by simply assuming $w \in \mathcal{L}$.

To solve the spy’s dilemma we use an interactive zero-knowledge proof, choosing \mathcal{L} , w and a “secret” concerning w . An authentic P will transmit parts of the secret so that V accepts $w \in \mathcal{L}$, but the transmission is otherwise useless to eavesdroppers and bogus verifiers. An imposter P^* would, with high probability, cause V to reject w .

Let \mathcal{L}_3 be “the set of strings that represent 3-colorable graphs” under some fixed graph-representation scheme. A graph is 3-colorable if its vertices can be assigned colors such that no adjacent vertices have the same color and no more than 3 colors are used. Let the set of colors be denoted \mathcal{C} . Both P and V have access to an *encryption function* $f: (\mathcal{C} \times \mathcal{R}) \rightarrow \mathcal{C}^e$, where \mathcal{R} contains long strings of h and t values (representing long sequences of coin tosses) and \mathcal{C}^e is a set of encrypted colors.

The common input string w_G represents a particular 3-colorable graph G of n vertices and m edges. The secret, known only to P , is a correct 3-coloring of G ; let c_i denote the color of vertex i under this coloring. An interactive zero-knowledge proof that $w_G \in \mathcal{L}_3$ is sketched below.

1. P applies a random permutation π to the colors: now each vertex i has color $\pi(c_i)$. Next, for each $i = 1 \dots n$, the prover forms a random string r_i from several coin tosses and computes $c_i^e = f(\pi(c_i), r_i)$. The encrypted vertex colors c_i^e are sent to V .
2. V saves $c_1^e \dots c_n^e$ and then chooses two adjacent vertices x and y at random and sends them to P .
3. P checks that (x, y) is really an edge in G . If not, the prover stops, having detected an imposter V that doesn’t know the protocol. If (x, y) is an edge, the prover sends the colors $\pi(c_x)$ and $\pi(c_y)$ and the values r_x and r_y to V .
4. V computes $c'_x = f(\pi(c_x), r_x)$ and $c'_y = f(\pi(c_y), r_y)$ and looks for inconsistencies with the transmission in Step 1, checking that $c'_x = c_x^e$ and $c'_y = c_y^e$. The verifier also looks for violations of 3-colorability, checking that $\pi(c_x), \pi(c_y) \in \mathcal{C}$ and that $\pi(c_x) \neq \pi(c_y)$. If any one of these checks fails, then V stops and rejects.
5. If the checks all pass, then P and V begin again in Step 1. If m^2 iterations of this protocol are completed without rejection, then V halts and accepts w_G .

Certainly the above protocol represents an interactive proof system for \mathcal{L}_3 . If $w_G \in \mathcal{L}_3$ then V accepts with probability 1 after m^2 iterations. If $w_G \notin \mathcal{L}_3$ then the prover must send an invalid coloring (one with adjacent vertices the same color or one that uses more than 3 colors) in Step 1, which will be detected in Step 4 with probability at least $1/m$ at each iteration. The probability that V halts and rejects after m^2 random probes is at least $1 - (1 - 1/m)^{m^2}$. A little calculation shows that this probability is sufficient (with the reasonable assumptions that $|w| \leq 2m \log_2 n$ and there are no isolated vertices in G).

Indeed, for an interactive proof it would be sufficient for P simply to send the vertex colors to V ; the extra steps are needed to ensure zero-knowledge. In Step 4, V “learns” that vertices x and y have different colors, which is no more than it

would learn from simply assuming $w_G \in \mathcal{L}_3$. The verifier gains no additional knowledge over time because the colors are randomly permuted and encrypted at each iteration. Even after several probes V has no idea how to 3-color the graph. A formal proof of zero-knowledge is rather too long to go into here: see Goldreich et al. [4] [5] for details.

BUT WILL IT WORK? Some nagging details must be addressed if this protocol is to be of any use to our spy. First, the proof of zero-knowledge depends upon an assumption that $f(\cdot, \cdot)$ is a *secure encryption* scheme, in the sense that it is not feasible to decrypt by deriving each $\pi(c_i)$ from c_i^e . Such a function is not known to exist.

Second, we must be assured that only the legitimate prover P could know a correct 3-coloring of G . The problem of finding 3-colorings of arbitrary graphs is known to be *NP-Hard*: as a consequence it is widely believed (but not proven) that any algorithm for finding 3-colorings must use exponential time on some graphs. Suppose we could build G such that the coloring is known by construction, but finding it independently requires time exponential in the size of G . Then we could secretly tell the coloring to P and (by choosing G large enough) be assured that any impostor computer would need, say, 10,000 years to find a 3-coloring. We also would have settled the most important open question in complexity theory today by proving that a famous set of problems known as \mathcal{NP} is not equivalent to another set called \mathcal{P} .

So our zero-knowledge scheme is not provably secure. Even without this assurance, however, the method is efficient and reliable enough to have been applied in practice. There are several encryption functions that have not been broken. A well-known one, for example, encrypts by multiplying large primes and relies on the fact that there is no known way to factor large numbers efficiently. And we can take other steps to reduce the chance of compromising the protocol: there exist zero-knowledge proofs that do not require encryption functions at all, and there exist problems that are “harder” than 3-coloring to solve.

FURTHER READING. The notions of interactive proof systems and knowledge complexity of proofs were first developed by Goldwasser et al. [6], [7]. Blum et al. [1] have since shown that zero-knowledge does not require interaction: that is, any interactive zero-knowledge proof can be replaced by one in which P sends messages to V but never receives any.

Goldreich et al. [4], [5] give several examples of zero-knowledge proofs, including the 3-colorability problem shown here, and discuss issues relating to secure protocols. Landau [8] describes what can happen when mathematicians and theoretical computer scientists get involved with problems of interest to the Department of Defense.

Technically, the zero-knowledge proof we’ve seen does reveal one bit of knowledge, namely that $w \in \mathcal{L}$. Fiege, Fiat and Shamir [3] have exhibited proofs that are truly zero-knowledge: the prover proves that he knows whether or not $w \in \mathcal{L}$, but doesn’t even reveal that fact. Perhaps someday we can extend zero-knowledge protocols to achieve a complete standstill of mathematical progress such as that attempted during the Renaissance. For example, maybe I could

demonstrate knowledge of the status of Fermat's last theorem¹ without revealing the proof or even the truth or falsehood of the statement. Heaven forbid.

REFERENCES

1. M. Blum, P. Feldman, and S. Micali, Non-interactive zero-knowledge and its applications, *Proceedings of the 29th Symposium on Foundations of Computer Science*, 1988, pp. 103–112.
2. C. B. Boyer, *A History of Mathematics*, Wiley and Sons, 1968.
3. U. Feige, A. Fiat, and A. Shamir, Zero knowledge proofs of identity, *Proceedings of the 19th Symposium on Theory of Computing*, 1987, pp. 210–217.
4. O. Goldreich, S. Micali, and A. Wigderson, Proofs that yield nothing but their validity and a methodology of cryptographic protocol design, *Proceedings of the 27th Symposium on Foundations of Computer Science*, 1986, pp. 174–187.
5. O. Goldreich, S. Micali, and A. Wigderson, Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proofs, *Journal of the Association for Computing Machinery*, (38) 1, 1991, pp. 691–729.
6. S. Goldwasser, S. Micali, and C. Rackoff, The knowledge complexity of interactive proof systems, *Proceedings of the 17th Symposium on Theory of Computing*, 1985, pp. 291–304.
7. S. Goldwasser, S. Micali, and C. Rackoff, The knowledge complexity of interactive proof systems, *SIAM J. Computing* (18) 1, February 1989, pp. 186–208.
8. S. Landau, Zero knowledge and the Department of Defense, *Notices of the AMS* (35) 1, January 1988, pp. 5–12.

Department of Mathematics and Computer Science
P.O. Box 2239
Amherst College
Amherst, MA 01002
ccm@cs.amherst.edu

PROBLEMS FOR SOLUTION

E. 36 *Proposed by B. H. Brown,*
Dartmouth College.

Show that the thirteenth of the month
is more likely to be Friday than any
one of the other days of the week.

—*American Mathematical Monthly*
40, (1933) p. 295

¹Please substitute “Goldbach’s Conjecture” for “Fermat’s Last Theorem” here.

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before February 28, 1994 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10322. *Proposed by Jiang Huanxin, student, FuDan University, ShangHai, China*

Let $ABCD$ and $AEFG$ be squares with the common vertex A and different edge lengths. Let $\theta = \angle EAD$ ($0 < \theta < \pi/2$). Suppose that EF and CD intersect at the point P . For which value of θ will AP be perpendicular to CF ?

10323. *Proposed by David E. Penney and Carl Pomerance, University of Georgia, Athens, GA.*

For a natural number n , let $t(n)$ be the sum of the divisors d of n in the range $1 \leq d < n$ with n/d being squarefree. Is there an integer n for which the sequence $n, t(n), t(t(n)), \dots$ is unbounded?

10324. *Proposed by William P. Wardlaw, United States Naval Academy, Annapolis, MD.*

Let a and m be positive integers and define the sequence $\langle x_n \rangle$ by $x_0 = 1$ and $x_{n+1} = a^{x_n}$. Show that there is a positive integer N such that $x_h \equiv x_k \pmod{m}$ whenever $N \leq h \leq k$.

10325. Proposed by Broderick Oluyede, Georgia State University, Atlanta, GA.

For $i = 1, 2, \dots, r$ and $j = 1, 2, \dots, c$, let $p_{i,j} \geq 0$, and assume that $\sum_{i=1}^r \sum_{j=1}^c p_{i,j} = 1$. Define $p_{i,\cdot} = \sum_{j=1}^c p_{i,j}$ and $p_{\cdot,j} = \sum_{i=1}^r p_{i,j}$. In addition, suppose that

$$p_{i+1,j+1} \sum_{h=1}^i \sum_{k=1}^j p_{h,k} \geq \sum_{h=1}^i p_{h,j+1} \sum_{k=1}^j p_{i+1,k}$$

for $0 < i < r$ and $0 < j < c$. Prove that

$$\sum_{h=1}^i \sum_{k=1}^j p_{h,k} \geq \sum_{h=1}^i p_{h,\cdot} \sum_{k=1}^j p_{\cdot,k}$$

for $0 < i < r$ and $0 < j < c$.

10326. Proposed by Ira Gessel, Brandeis University, Waltham, MA.

For r a positive integer, let K_r be the smallest positive integer such that

$$\frac{K_r}{n+r} \binom{2n}{n}$$

is an integer for all $n \geq 0$. Show that

$$K_r = \frac{r}{2} \binom{2r}{r}.$$

10327. Proposed by Jerome Minkus, Berkeley, CA.

Find the simple continued fraction for $(e+3)/4$.

10328. Proposed by A. Keith Austin, The University of Sheffield, Sheffield, England.

Let A and B be sets such that $A \cap B = \emptyset$ and $A \cup B$ is the unit square $[0, 1] \times [0, 1]$. Prove or disprove the following:

(a)* Either there is a continuous function $f: [0, 1] \rightarrow A$ with $f(0) = (0, y_0)$ for some y_0 and $f(1) = (1, y_1)$ for some y_1 , or there is a continuous function $g: [0, 1] \rightarrow B$ with $g(0) = (x_0, 0)$ for some x_0 and $g(1) = (x_1, 1)$ for some x_1 .

(b) f and g as in part a cannot both exist.

10329. Proposed by Gérard Letac, Université Paul Sabatier, Toulouse, France.

Let $f(x)$ is a positive continuous function defined for $0 < x < 1$ such that, for all u with $0 < u < 1$, one has $\int_u^1 f(x)f(u/x) dx = u^{1/2}$. Prove that

$$f(x) = \sqrt{\frac{2x}{\pi(1-x^2)}}.$$

NOTES

Notes: (10323) A related sequence, called the *aliquot sequence* of n is generated by using a function $s(n)$ which is the sum of all divisors d of n in the interval $1 \leq d < n$. Some examples of aliquot sequences are: 9, 4, 3, 1, 0, 0, ...; 6, 6, ...; and 220, 284, 220, ... It is unknown whether all aliquot sequences are eventually periodic; the case of $n = 276$ is unresolved at this time. **(10327)** The standard reference for continued fractions is O. Perron, *Die Lehre von den Kettenbrüchen*. The fourth chapter describes the continued fraction for e and related “Hurwitz continued fractions”.

SOLUTIONS

Uniqueness from Asymptotic Behavior

E 3449 [1991, 553]. *Proposed by Mark A. Pinsky, Northwestern University, Evanston, IL.*

Suppose s is a continuous real-valued function on $[0, +\infty)$ such that s is differentiable on $(0, +\infty)$, $0 \leq s(t) \leq t^2$, and

$$\frac{ds}{dt} = t + \sqrt{t^2 - s} \quad (t > 0).$$

Prove that s is unique and obtain a closed formula for s .

Solution I by J. B. Thoo, student, University of California, Davis, CA. By direct substitution it is easily verified that $s(t) = \frac{3}{4}t^2$ satisfies the requirements.

To establish that this $s(t)$ is the unique solution, we will show that any two solutions $s_1(t)$ and $s_2(t)$ must be identical. Let us define $g(t) = (s_1(t) - s_2(t))^2$, which is clearly non-negative. Then for all $t > 0$,

$$\begin{aligned} g'(t) &= 2(s_1(t) - s_2(t))(s_1'(t) - s_2'(t)) \\ &= -2 \frac{(s_1(t) - s_2(t))^2}{(t^2 - s_1(t))^{1/2} + (t^2 - s_2(t))^{1/2}} \\ &\leq 0. \end{aligned}$$

Hence, $g(t) \leq g(0)$ for all $t > 0$. But since $0 \leq s(t) \leq t^2$ implies $s_1(0) = 0$ and $s_2(0) = 0$, then also $g(0) = 0$; hence, for all $t > 0$, $g(t) \leq 0$. Since g is both a non-negative function and, it now appears, a non-positive one as well, it must be identically zero, and so therefore $s_1 = s_2$, as claimed.

Solution II by Frédéric Brulois, California State University–Dominguez Hills, Carson, CA. Re-write the given condition in the form $s'(t) - t = (t^2 - s(t))^{1/2}$. Square it and differentiate it to obtain $2(s'(t) - t)s''(t) = s'(t)$. This is a first-order

homogeneous equation in $s'(t)$, which can be solved by standard techniques to obtain $s'^2(t)(3t - 2s'(t)) = C$. Thus, using the parameter $p = s'(t)$, we get $t = (2/3)p + (C/3)p^{-2}$ and $s = (1/3)p^2 + (2C/3)p^{-1}$. Since $s'(t)$ lies between t and $2t$ for all $t > 0$, the only possible value of C that would permit this equation to hold for arbitrarily small t is 0. It follows from the parametric solution that $t = 2p/3$ and $s = p^2/3$. Thus $s(t) = 3t^2/4$.

Solution III by Kiran S. Kedlaya, student, Harvard University, Cambridge, MA. Define $f(x) = \frac{1}{2}(1 + (1 - x)^{1/2})$, and note that $x = 3/4$ is a fixed point of this function. Also note that $|f'(x)| < 7/8$ for $1/2 < x < 7/8$ and that this interval is taken into itself by f . Furthermore, any sequence defined inductively by $x_{n+1} = f(x_n)$, with $x_0 \in [0, 1]$, eventually enters this attracting basin and converges to $3/4$. In particular, we choose $x_0 = 0$.

We prove by induction that $x_{2n}t^2 \leq s(t) \leq x_{2n+1}t^2$ for all t and all n . From this and the fact that $x_n \rightarrow 3/4$, we may conclude that $\frac{3}{4}t^2 \leq s(t) \leq \frac{3}{4}t^2$, and thence that $s(t) = \frac{3}{4}t^2$.

The statement with $n = 0$ is hypothesized, so let us show that if $x_{2k}t^2 \leq s(t) \leq x_{2k+1}t^2$ holds for some $k \geq 0$, then $x_{2k+2}t^2 \leq s(t) \leq x_{2k+3}t^2$ holds also. We have

$$2tx_{2k+2} = t + \sqrt{t^2 - x_{2k+1}t^2} \leq t + \sqrt{t^2 - s} \leq t + \sqrt{t^2 - x_{2k}t^2} = 2tx_{2k+1}$$

where the equalities follow from the inductive definition of x_n and the inequalities follow from the induction hypothesis. Then by integrating, we obtain

$$\int_0^t 2tx_{2k+2} dt \leq \int_0^t s'(t) dt \leq \int_0^t 2tx_{2k+1} dt$$

$$x_{2k+2}t^2 \leq s(t) \leq x_{2k+1}t^2,$$

where we have used the fact that $s(0) = 0$. Then, by similar reasoning, we reach the desired conclusion that

$$x_{2k+2}t^2 \leq s(t) \leq x_{2k+3}t^2.$$

Editorial comment. These three solutions are representatives of the principal methods of solution. These may be summarized as follows.

Method I: Guess the answer. Prove that it works. Then give careful attention to proving uniqueness.

Method II: Transform the differential equation by a change of variable or further differentiation into an equation whose complete solution can be found by standard methods. Then impose the restriction that $0 \leq s(t) \leq t^2$.

Method III: Use the differential equation to iteratively produce explicit refinements of the requirement that $0 \leq s(t) \leq t^2$ for all $t > 0$. Existence and uniqueness will then follow from general fixed-point arguments. In this method, a familiar method of proof of the existence and uniqueness theorem of differential equations is applied, exploiting a global inequality on the solution to control $\int_0^t s'(t) dt$.

Solved by 55 readers, some submitting more than one solution, and the proposer. This yielded 21 solutions by Method I, 19 by Method II, 16 by Method III, and 4 hybrids. In addition there were 7 submissions found to be incomplete or inaccurate.

Graceful Permutations

E 3455 [1991, 646]. *Proposed by D. G. Rogers, University of Aberdeen, Scotland, and Howard University, Washington, DC.*

It is known that if $n \equiv 0 \pmod{4}$ or $n \equiv 1 \pmod{4}$, then there exist permutations (x_1, x_2, \dots, x_n) of $(1, 2, \dots, n)$ such that the differences $|x_k - k|$, $1 \leq k \leq n$, are all distinct. (Cf. E 3269 [1988, 554; 1989, 843].) Prove that the number of such permutations is a multiple of 4.

Solution by M. Roth and O. Šuch, Queen's University, Kingston, Ontario, Canada. Let S be the set of permutations such that the specified differences are all distinct. Let $n > 1$ to assure that the identity does not belong to S . We show that the number of such permutations is a multiple of 4 by defining two involutions π and ρ on S such that $\pi\rho(\sigma) = \rho\pi(\sigma)$ and $\pi(\sigma) \neq \rho(\sigma)$ for any $\sigma \in S$. If we also show that π and ρ do not fix any element of S then the action of these operations splits S into disjoint orbits of size 4, which proves $|S| \equiv 0 \pmod{4}$.

Letting $\sigma = x_1, \dots, x_n$, $\pi(\sigma) = y_1, \dots, y_n$ and $\rho(\sigma) = z_1, \dots, z_n$, we define $\pi(\sigma)$ and $\rho(\sigma)$ explicitly by $y_k = j$ if and only if $x_j = k$, and $z_k = j$ if and only if $x_{n+1-k} = n+1-j$. By construction, these produce permutations, preserve the set of differences, and are involutions. Note that π takes a permutation to its inverse. An element of S cannot interchange a pair of points, and can have at most one fixed point, so π fixes no element of S . If $\rho(\sigma) = \sigma$, then $x_k = j$ if and only if $x_{n+1-k} = n+1-j$, in which case the differences for positions k and $n+1-k$ have the same magnitude and $\sigma \notin S$.

By direct calculation, both $\pi\rho(\sigma)$ and $\rho\pi(\sigma)$ have $n+1-k$ in position $n+1-j$ if and only if $x_k = j$, so $\pi\rho = \rho\pi$. All that remains to be shown is that these maps are not equal on any member of S . Note that $\rho(\sigma)$ has $n+1-x_{n+1-k}$ in position k . If $\pi(\sigma) = \rho(\sigma)$, then σ also has k in position $n+1-x_{n+1-k}$. This makes the absolute difference between a position and its value the same at position $n+1-k$ and position $n+1-x_{n+1-k}$. If $\sigma \in S$, the differences are distinct for distinct positions, and hence $k = x_{n+1-k}$ for all k . This is satisfied only by permutation which is not in S .

Editorial comment. Only a few solvers made explicit mention of the fact that the statement of the problem needed to be modified to require $n > 1$. The term “graceful” for the permutations with this property was suggested by Albert Nijenhuis.

Solved also by D. Callan, R. J. Chapman (U.K.), P. Čížek (student, Czech Republic), M. Dindos (Slovakia), J. Fukuta (Japan), L. L. Gardner, R. High, A. A. Jagers (The Netherlands), I. Kastanas, K. S. Kedlaya (student), O. P. Lossers (The Netherlands), J. H. Nieto (Venezuela), A. Nijenhuis, J. H. Steelman, C. Voas, National Security Agency Problems Group, Shreveport Problem Solving Group (LSU), and the proposer.

A Variant of the Erdős-Faber-Lovász Conjecture

6664 [1991, 655]. *Proposed by Paul Erdős, Hungarian Academy of Sciences, Budapest*

Let G be a graph whose edges can be covered by n complete subgraphs with n vertices each (i.e., G is the union of n copies of K_n , with no restrictions on shared vertices).

(a) Prove that the chromatic number of G is less than $1 + n\sqrt{n-1}$.

(b) Prove that this bound is asymptotically best possible, i.e., if $f(n)$ is the maximum chromatic number of a graph constructed in this way, then $f(n) = \{1 + o(1)\}n^{3/2}$.

Solution of (a) by Ilias Kastanas, California State University, Los Angeles, CA, and by Charles Vanden Eynden, Illinois State University, Normal, IL (independently). A graph with chromatic number k has at least $\binom{k}{2}$ edges, because if there is no edge between the set of vertices of color i and the set of vertices of color j , then colors i and j can be combined into a single color. On the other hand, G has at most $n\binom{n}{2}$ edges, and the inequality $k(k-1) \leq n^2(n-1)$ implies $k < 1 + n\sqrt{n-1}$.

Solution of (b) by Richard Holzsager, American University, Washington, DC. Consider the affine plane of order p , where p is a prime. There are p^2 points and $p^2 + p$ lines of size p , such that each pair of points appears in a unique line. If we view these lines as cliques (complete graphs) on the points, then we have expressed the complete graph K_{p^2} as a union of $p^2 + p$ cliques of size p . We now expand each point into a clique of $p + 1$ points to express $K_{p^3+p^2}$ as a union of $p^2 + p$ cliques of size $p^2 + p$. Hence $f(p^2 + p) \geq p^3 + p^2 = (1 + o(1))(p^2 + p)^{3/2}$.

Now, let n be arbitrary, and fix $\varepsilon > 0$. By the prime number theorem, the number of primes below x is eventually greater than the number of primes less than $(1 - \varepsilon)x$, meaning there are primes between $(1 - \varepsilon)x$ and x , if x is large enough. Taking n large enough and $x = \sqrt{n + 1/4} - 1/2$, we obtain a prime p with $(1 - 2\varepsilon)n < p(p + 1) < n$. Therefore, $f(n) \geq f(p^2 + p) \geq (1 + o(1))(p^2 + p)^{3/2} = (1 + o(1))n^{3/2}$.

Editorial comment. This problem was first received from the proposer by the editors in 1987. In 1988, P. Horák heard of the problem and found a solution, which was published as “A coloring problem related to the Erdős-Faber-Lovász conjecture,” *J. Combinatorial Theory*, Ser. B 50 (1990), 321–322.

Suppose we add the constraint that each edge of G appears in exactly one clique (note that this is violated by the construction in (b)). The Erdős-Faber-Lovász conjecture is that in this case the chromatic number is exactly n .

Zoltan Füredi improved the upper bound of (a) when n is of the form $q^2 + q$. He proved that $q^3 + q^2$ is an upper bound in this case, which makes the projective plane construction of (b) optimal when $n = q^2 + q$ and q is a prime power.

The problem was completely solved by all three solvers cited above. Solutions were also given by the proposer and by Z. Füredi.

More Pigeons on the Circle

E 3463 [1991, 852]. *Proposed by Donald E. Knuth, Stanford University, Stanford, CA.*

Let S be a set of m distinct points on the unit circle such that no two are diametrically opposite. For a fixed integer $n \leq m/2$, suppose that we mark every point p in S such that fewer than n of the remaining points in S lie in the semicircle counterclockwise from p . Prove that at most n points are marked.

Solution by John H. Lindsey II, Fort Myers, FL. Fix n and delete unmarked points one by one, each time allowing unmarked points to become marked as the semicircles empty out, until only $2n$ points remain or all remaining points are marked. At this time a point is marked only if the n th point later is more than a

semicircle away. If that point is also marked, then traversing $2n$ points travels more than the full circle, which happens only if fewer than $2n$ points remain. Hence the stopping condition occurs when exactly $2n$ points remain, and the marked points consist of exactly one point from each pair of points separated by $n - 1$ points in each direction. This implies there were at most n marked points in the original set.

Solved by 26 other readers and the proposer.

A Half Step Towards Carmichael's Conjecture

6671 [1991, 862]. *Proposed by Carl Pomerance, University of Georgia, Athens, GA.*

Let $V(x)$ denote the number of distinct values not exceeding x taken on by Euler's arithmetical function ϕ . Let $V^*(x)$ denote the number of these values with a unique pre-image. For example, $V(15) = 7$, $V^*(15) = 0$.

R. D. Carmichael conjectured that $V^*(x) = 0$ for all x . Prove the weaker assertion that $\liminf_{x \rightarrow \infty} \{V^*(x)/V(x)\} < 1$.

Solution by L. E. Mattics, University of South Alabama, Mobile, AL. Let $\varepsilon = 2^{1/2} - 1$ and let $V_0(x)$ be the number of values of $\phi(w)$ not exceeding x such that w is odd. We will show that we can prove that $\liminf_{x \rightarrow \infty} V^*(x)/V(x) \leq 2^{-1/2}$ regardless of whether or not the following proposition holds.

Proposition. *For every positive integer N there is an $x \geq N$ such that $V_0(x/2) > \varepsilon V(x)$.*

If the proposition does hold then there are arbitrarily large x such that

$$V\left(\frac{x}{2}\right) \geq V_0\left(\frac{x}{2}\right) + V^*\left(\frac{x}{2}\right) \geq \varepsilon V(x) + V^*\left(\frac{x}{2}\right) \geq (\varepsilon + 1)V^*\left(\frac{x}{2}\right) \geq 2^{1/2}V^*\left(\frac{x}{2}\right)$$

so $\liminf_{x \rightarrow \infty} V^*(x)/V(x) \leq 2^{-1/2}$.

Assume from now on that the proposition does not hold. Then there is an integer N such that for all $x \geq N$, $V_0(x/2) \leq \varepsilon V(x)$. If $(2, u) = 1$ and $\phi(2^a u) \leq x$ has only one pre-image, then $a \geq 2$ and $\phi(u) \leq x/2^{a-1}$; and if v is odd and $\phi(u) = \phi(v)$, then $\phi(2^a u) = \phi(2^a v)$, so $u = v$. This implies that $V^*(x) \leq \sum_{a=2}^{\infty} V_0(x/2^{a-1})$ where $V_0(c) = 0$ if $c < 1$.

Now let $m = \lfloor \log_2(x/N) \rfloor + 1$ then

$$\begin{aligned} V^*(x) &\leq \sum_{a=2}^m \varepsilon^{a-1} V(x) + \sum_{a=m+1}^{\infty} V_0\left(\frac{x}{2^{a-1}}\right) \\ &\leq \frac{\varepsilon}{1-\varepsilon} V(x) + \left(V_0(N) + V_0\left(\frac{N}{2}\right) + \cdots \right). \end{aligned}$$

Since $V(x) \rightarrow \infty$ as $x \rightarrow \infty$ and N is fixed we have $\liminf_{x \rightarrow \infty} V^*(x)/V(x) \leq 2^{-1/2}$.

Editorial comment. All solutions provided an upper bound on the quantity $\liminf_{x \rightarrow \infty} \{V^*(x)/V(x)\}$. The best value obtained to date was $1/2$, which was given by the proposer. All solutions considered the set $\Phi_o = \{m: m = \phi(2k+1) \text{ for some } k\}$, and noted that, if m is $\phi(n)$ for some n , then there is an integer h

with $m/2^h \in \Phi_o$. This should be compared to problem E 3661 [1990, 63; 1991, 443] in which examples of $\phi(n) \notin \Phi_o$ were given.

Solved also by I. Kastanas and the proposer.

Consecutive Convergents

10187 [1992, 60]. *Proposed by Irving Adler, North Bennington, VT.*

Suppose n_{k-1}/d_{k-1} and n_k/d_k are consecutive convergents of the simple continued fraction for some real number α in $(0, 1)$. Assume you are given only the values of d_{k-1} and d_k . Construct an algorithm for determining the values of n_{k-1} and n_k .

Solution by Nicholas C. Singer, Annandale, VA. The problem as stated has two possible solutions, because the convergents to α and $1 - \alpha$ have essentially the same sequence of denominators. That is, if $0 < \alpha < 1/2$, then α has the continued fraction expansion $\alpha = [0, a_1, a_2, a_3, \dots]$ with $a_1 \geq 2$; and then $1 - \alpha = [0, 1, a_1 - 1, a_2, a_3, \dots]$. Using the standard recurrence relation

$$d_k(\alpha) = a_k d_{k-1}(\alpha) + d_{k-2}(\alpha), \quad d_{-2}(\alpha) = 1, d_{-1}(\alpha) = 0,$$

we conclude that for $k \geq 1$, $d_k(1 - \alpha) = d_{k-1}(\alpha)$. We need exactly one bit of additional information to get a unique answer: (i) is α greater than or less than $1/2$? or (ii) what is the parity of k ?

It is immediate, using the recurrence relations, that $d_k/d_{k-1} = [a_k, a_{k-1}, a_{k-2}, \dots, a_1]$. The quotients and convergents are calculated using the usual continued fraction (which is equivalent to the Euclidean algorithm). The n_k satisfy the same recurrence as the d_k with the initial conditions replaced by $n_{-2} = 0$, $n_{-1} = 1$. In addition, $n_0 = a_0 = 0$ so we also have $n_k/n_{k-1} = [a_k, a_{k-1}, a_{k-2}, \dots, a_2]$. (The case $k = 1$ is special since $n_0 = 0$.) That is, we take n_k and n_{k-1} to be the numerator and denominator of the penultimate convergent of the continued fraction expansion of d_k/d_{k-1} .

The usual application of the Euclidean algorithm always gives $a_1 \geq 2$. However, $[a_k, a_{k-1}, a_{k-2}, \dots, a_2, a_1] = [a_k, a_{k-1}, a_{k-2}, \dots, a_2, a_1 - 1, 1]$, which leads to the alternative expansion $n_k/n_{k-1} = [a_k, a_{k-1}, a_{k-2}, \dots, a_2, a_1 - 1]$. This corresponds to $1 - \alpha = [0, 1, a_1 - 1, a_2, \dots, a_k, \dots] > 1/2$. This expansion of n_k/n_{k-1} has k convergents, whereas the previous expansion had $k - 1$. Hence knowing the answer to either question (i) or question (ii) allows us to produce the unique correct result.

Solved also by D. Callan, R. J. Chapman (U. K.), D. Chinitz (student), C. H. Ebersole, B. Haible (Germany), R. J. Hendel, R. High, O. P. Lossers (The Netherlands), A. Nijenhuis, J. H. Steelman, R. Stong, B. M. M. de Weger (The Netherlands), E. A. Weinstein, O. Wyler, Anchorage Math Solutions Group, National Security Agency Problems Group, and the proposer.

Complex Conjugation of $\mathbb{C}(z)$

10191 [1992, 61]. *Proposed by Dragomir Ž. Đoković, University of Waterloo, Waterloo, Ontario, Canada.*

Let G be the group of \mathbb{C} -automorphisms of the function field $\mathbb{C}(z)$ and Σ the set of involutory automorphisms of $\mathbb{C}(z)$ which extend the complex conjugation on

C. Show that Σ splits into two orbits under the action $G \times \Sigma \rightarrow \Sigma$, $(\alpha, \beta) \mapsto \alpha \circ \beta \circ \alpha^{-1}$. (Thus there are only two essentially different ways of extending the complex conjugation to an involutory automorphism of $\mathbb{C}(z)$.)

Solution by Robin J. Chapman, University of Exeter, Exeter, U.K. It is well known that if $\alpha \in G$ then α is determined by $\alpha(z) = (az + b)/(cz + d)$ where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a non-singular matrix over \mathbb{C} . Two different choices of A give the same α if and only if they are scalar multiples of each other. Also, composition in G corresponds to matrix multiplication. Furthermore, an automorphism β of $\mathbb{C}(z)$, restricting to conjugation on \mathbb{C} , is also determined by $\beta(z) = (pz + q)/(rz + s)$ where $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ is a non-singular matrix over \mathbb{C} ; and again, two different B give the same β if and only if they are scalar multiples. Now we easily compute that $\beta^2(z) = (tz + u)/(vz + w)$ where $C = \begin{pmatrix} t & u \\ v & w \end{pmatrix} = \bar{B}B$. Hence $\beta \in \Sigma$ if and only if $\bar{B}B = \lambda I$ for some $\lambda \in \mathbb{C}$. If $\beta \in \Sigma$ then $\lambda^2 = \det \bar{B} \det B = |\det B|^2$. Hence $\lambda = \pm |\det B|$. As replacing B by μB changes λ to $|\mu|^2 \lambda$ and $|\det B|$ to $|\mu|^2 |\det B|$ then the sign of λ is an invariant of β . If $B = \begin{pmatrix} 0 & 1 \\ \pm 1 & 0 \end{pmatrix}$ then $\bar{B}B = \pm I$ and so both signs occur. Call β positive or negative according to the sign of λ .

I claim now that if $\alpha \in G$ and $\beta \in \Sigma$ then $\alpha \circ \beta \circ \alpha^{-1}$ is positive if and only if β is. If α and β are represented by the matrices A and B respectively then $\alpha \circ \beta \circ \alpha^{-1}$ is represented by $B' = \bar{A}^{-1}BA$. Now $\bar{B}'B' = A^{-1}\bar{B}A\bar{A}^{-1}BA = A^{-1}\bar{B}BA = \lambda I$ and the claim follows. Hence Σ splits into at least two G -orbits.

Take $\beta \in \Sigma$. We may represent β by a matrix B with $\det B = -\text{sgn}(\beta)$. Hence $\bar{B} = -(\det B)B^{-1}$ and $B = \begin{pmatrix} a & b \\ c & -\bar{a} \end{pmatrix}$ where $a \in \mathbb{C}$, $b \in \mathbb{R}$, $c \in \mathbb{R}$ and $|a|^2 + bc = \pm 1$. Now if $|\zeta| = 1$ and $A_1 = \begin{pmatrix} \zeta & 0 \\ 0 & \bar{\zeta} \end{pmatrix}$ then $B' = \bar{A}_1^{-1}BA_1 = \begin{pmatrix} \zeta^2 a & b \\ c & -\zeta^2 \bar{a} \end{pmatrix}$ so that by a suitable choice of ζ we may assume that $B' = \begin{pmatrix} a' & b' \\ \zeta & -a' \end{pmatrix}$ where $a' \in \mathbb{R}$ and $\det B' = \pm 1$. Hence B' has characteristic polynomial $X^2 \pm 1$ and so there is a real matrix A_2 with $A_2^{-1}B'A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ if β is positive and $A_2^{-1}B'A_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ if β is negative. Hence if $A = A_1A_2$ then $\bar{A}^{-1}BA = \begin{pmatrix} 0 & 1 \\ \pm 1 & 0 \end{pmatrix}$ and so there are at most two G -orbits in Σ .

Editorial comment. Robin Chapman also provided a cohomological interpretation of the result. If we let $\Gamma = \text{Gal}(\mathbb{C}/\mathbb{R})$, then Γ acts on $G \cong \text{PGL}_2(\mathbb{C})$ by conjugation and it is not hard to see that the elements of Σ correspond to 1-cocycles of Γ in G and that two elements of Σ correspond to cohomologous cocycles if and only if they lie in the same G -orbit. Hence the set of orbits corresponds to the set $H^1(\Gamma, G)$. Using the exact sequence of Γ -modules

$$1 \rightarrow \mathbb{C}^* \rightarrow \text{GL}_2(\mathbb{C}) \rightarrow \text{PGL}_2(\mathbb{C}) \rightarrow 1$$

a standard theorem of Galois cohomology (Jean-Pierre Serre, *Local Fields*, Springer-Verlag, 1978, X.5), shows that the connecting map $\delta: H^1(\Gamma, G) \rightarrow H^2(\Gamma, \mathbb{C}^*)$ is an isomorphism. Now as Γ is cyclic, it is immediate that $H^2(\Gamma, \mathbb{C}^*) \cong \mathbb{R}^*/N(\mathbb{C}^*) \cong \{\pm 1\}$ where $N: \mathbb{C} \rightarrow \mathbb{R}$ is the norm map, and the result follows. More generally if we replace \mathbb{R} and \mathbb{C} by K and L where L/K is a quadratic extension with Galois group Γ then the corresponding result is that the G -orbits of Σ are in one-to-one correspondence with $H^1(\Gamma, G) \cong H^2(\Gamma, L^*) \cong K^*/N_{L/K}(L^*)$.

Now $H^2(\Gamma, L^*)$ is the relative Brauer group $\text{Br}(L/K)$, and as $[L:K] = 2$ this can be interpreted as the set of equivalence classes of 1-dimensional Severi-Brauer

varieties over K split by L . These are the projective curves defined over K which become isomorphic to the projective line after base change to L . If $\beta \in \Sigma$ then the fixed field $L(z)^\beta$ is the function field of the corresponding Severi-Brauer variety. If $L = \mathbb{C}$ and $K = \mathbb{R}$ and if β is positive, then $\mathbb{C}(z)^\beta \cong \mathbb{R}(t)$, the function field of the projective line over \mathbb{R} ; and if β is negative, then $\mathbb{C}(z)^\beta \cong \mathbb{R}(x, y | x^2 + y^2 + 1 = 0)$, the function field of the conic C with homogeneous equation $X_1^2 + X_2^2 + X_3^2 = 0$ which has no points defined over \mathbb{R} . Explicitly, if $\beta(z) = z$, then β is positive and $\mathbb{C}(z)^\beta = \mathbb{R}(z)$; while if $\beta(z) = -1/z$, then β is negative and $\mathbb{C}(z)^\beta = \mathbb{R}(x, y)$ where $x = (z - 1/z)/2$ and $y = i(z + 1/z)/2$ satisfy $x^2 + y^2 = -1$.

The proposer's proof that there are at most two orbits involved showing that any matrix B with $\overline{B}B = I$ can be written as $\overline{A}^{-1}A$, for which he referred to D. Ž. Đoković, "On some representations of matrices", *Linear and Multilinear Algebra*, 4 (1976), 33–40.

Solved also by D. Callan and the proposer.

Collaborating editors: *David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Jerrold R. Griggs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pflieger, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, and William E. Watkins.*

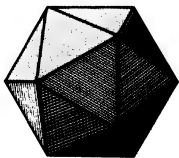
"You know, for a mathematician he did not have enough imagination. But he has become a poet and now he is doing fine....."

—Hilbert (to Cassirer,
about a former student)

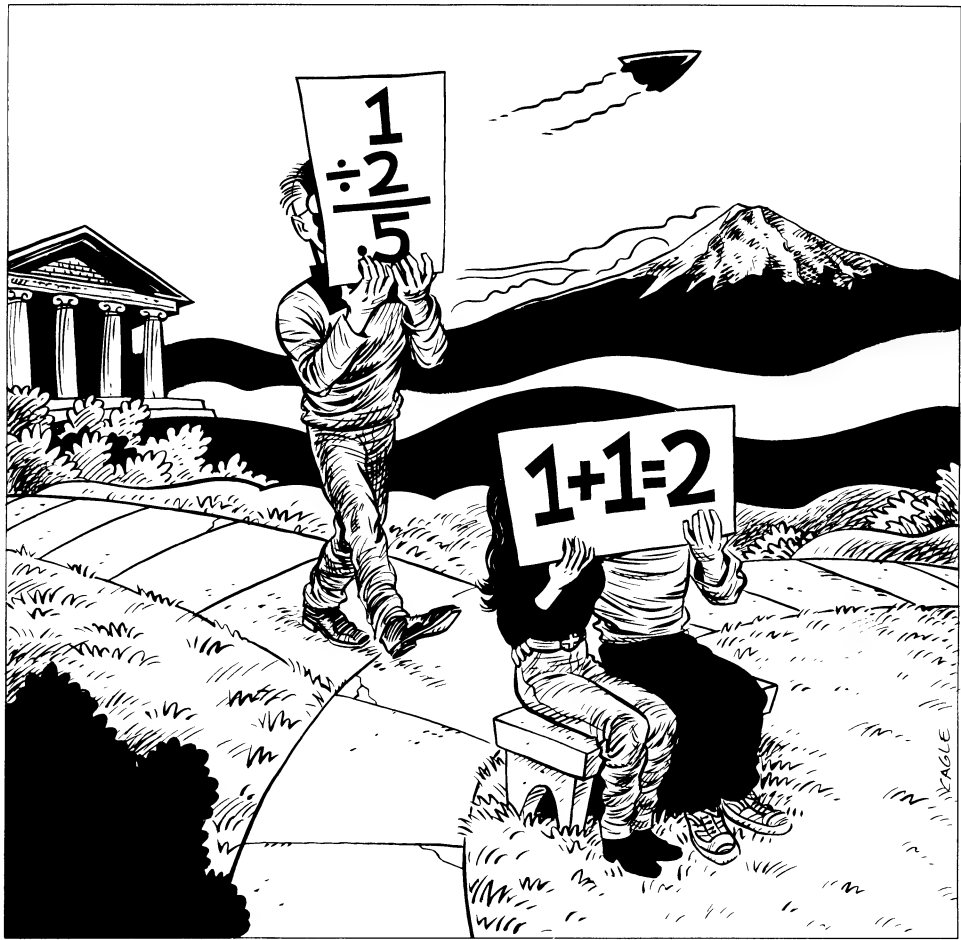
Answer to Picture Puzzle
(p. 661)
George Pólya.

Answer to Who Was the Author
(p. 681)
James Joseph Sylvester.

The American Mathematical Monthly



Volume 100, Number 8 / OCTOBER 1993



NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

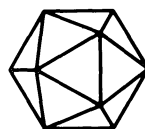
Microfilm Editions: University Microfilms International, Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 8 / OCTOBER 1993

(ISSN 0002-9890)



Contents

ARTICLES

Thomas Archer Hirst—Mathematician Xtravagant IV. Queenwood, France
and Italy / J. HELEN GARDNER and ROBIN J. WILSON 723

Thoughts on Innumeracy: Mathematics Versus the World? / PETER L.
RENTZ John Allen Paulos Replies 732

Reflections on Rippling Water / MICHEL MENDES FRANCE 743

The Principal Axis Theorem over Arbitrary Fields /
DAVID MORNHINWEG, DANIEL B. SHAPIRO,
and K. G. VALENTE 749

The Fifty-Third William Lowell Mathematical Competition /
LEONARD F. KLOSINSKI, GERALD L. ALEXANDERSON,
and LOREN C. LARSON 758

A Visual Explanation of Jensen's Inequality / TRISTAN NEEDHAM 768

The Index of a Constrained Critical Point / CATHERINE HASSELL
and ELMER REES 772

On Some Irrational Decimal Fractions / NORBERT HEGYVÁRI 779

FEATURES

COMMENTS 722

PICTURE PUZZLE 748

NOTES 781

LETTER TO THE EDITOR

UNSOLVED PROBLEMS 789

Open Problems in Pattern Avoidance / JAMES CURRIE 790

THE AUTHORS 794

PROBLEMS AND SOLUTIONS 796

REVIEWS

Ordinary Differential Equations. By Vladimir I. Arnol'd /
FRED BRAUER 810

TELEGRAPHIC REVIEWS 812

Thomas Archer Hirst— Mathematician Xtravagant IV. Queenwood, France and Italy

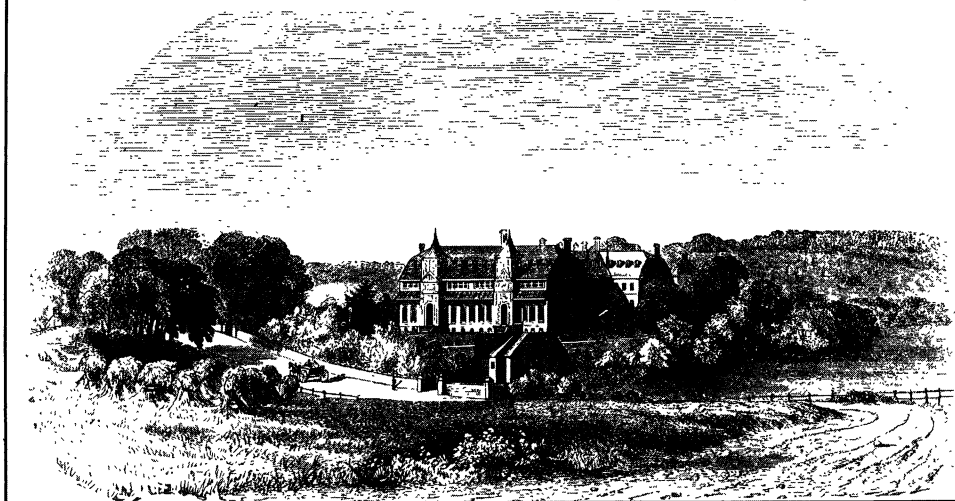
J. Helen Gardner and Robin J. Wilson

I occupied myself most of the day by sketching out a kind of inaugural lecture for Queenwood. It is now certain that in August I shall commence my life as Tutor there. It has for me its attractions and at the same time its onerous duties and responsibilities—I meet both cheerfully, and I hope for strength and courage to fulfil the task I have chosen for myself in the world.

Thomas Hirst returned to England in mid-summer 1853. At a time when there were limited job opportunities for a young mathematician, he was lucky to be offered a teaching post at Queenwood College, near Salisbury, where John Tyndall and his chemist friend Edward Frankland had taught before their Marburg days.

Queenwood College

Queenwood itself is a beautiful spot, it stands in a rich undulated chalk district, the small knowls and vallies are always graceful and smooth and the rich woods, with their beautiful beeches, yews and elms have a soothing effect. The building itself is interesting on many accounts, first, its architecture which is in a novel and picturesque style, mostly Italian. Secondly, its inward arrangements, which are the most convenient and beautiful that I have seen, and thirdly its associations, for this is the celebrated Harmony Hall, where the socialists first practically tried to live by the law of love, and of course miserably failed . . . now it makes one of the most beautiful schools I ever saw, and from all accounts the scholastic arrangements are just as good.



Queenwood was partly boys' elementary school, partly mechanics' institute—a sort of technical college. There was a strong Pestolozzian influence, in that the teaching emphasized practical work by the pupils. For example, Hirst taught geometry in the context of surveying, rather than as theorems from Euclid. This experience was to prove useful later when he emerged as a reformer of geometry teaching in schools.

14th August 1853: We have now got thoroughly to work. I have 13 hours a week teaching, and two lectures; and I get more and more to love my work. The profession of schoolmaster is no drudgery, but when properly undertaken a noble task, and a healthy discipline. Yes, I have come to the conclusion that I have found my proper task, and to the determination to fulfil it to the best of my ability. At present it occupies nearly all my time, and must do until I am thoroughly master of my best plan for tuition. That done, then I sit down to my own investigations.

He was also developing a reputation for giving public lectures on physics.

30th October 1853: ... I have now conquered to a great extent all nervousness—it would be no task to speak to a thousand people about a subject with which I was well acquainted. Nevertheless, I have not yet got the tact to know what parts will be best appreciated. I found afterwards that exactly the parts on which I had laid least stress were best appreciated, and vice versa. This talent which I lack is essentially necessary to a popular lecturer...

When he could, he escaped to London to see Tyndall and to attend popular science lectures at the Royal Institution, where Tyndall was now working.

21st January 1854: Friday evening I attended Faraday's lecture—"Electricity of Induction static and dynamic effects." The lecture room was filled with a very brilliant audience, and the lecture itself pleased me much... Perhaps the lecturer's manner, person and celebrity attracted me most. There was about him such a total absence of mannerism and pretension, such geniality and gentleness...

As John and I were sitting writing, after tea this evening, Faraday himself paid us a brief visit... The room smelt villainously of tobacco, although John hurriedly scattered some eau de Cologne on the carpet. The candles too just went out and Faraday made his entrance almost in the dark. After a short time, during which Faraday had bowed to me, John remarked that I was a friend of his. "Oh, indeed!" says Faraday, fetching a chair, "well, let us all sit down, and have a look at one another." He did sit down, and after looking at me for a minute got up and shook me very kindly by the hand, saying it was a pleasure to him to know any of Tyndall's friends...

Hirst's attitude to women was somewhat prudish and patronizing. While in Marburg, he had struck up an acquaintance with a young lady called Anna Martin.

3rd July 1854: ... Instinctively I got to admire her, her artlessness, her affection for her own family, her honest independence, and even waywardness towards me, and finally her frank friendship for me in spite of all my bluntness and scolding—all these things, no doubt, besides many others, drew her nearer to me...

They were married late in 1854 at Anna's home in County Down, and returned to Queenwood after spending their honeymoon in Paris. Hirst was blissfully happy.

18th February 1855: ... I have convinced myself that she is and will be a true and devoted companion. There is in her a far deeper devotedness than I could have anticipated... I have found that her happiness consists not in comforts and luxuries, but rests on the far higher and more womanly consciousness that she is necessary to her husband's happiness... Her failings, as failings of course she has, I can trace almost entirely to her irregular life and training... yet when I *do* see her bustling about in her own cheerful, merry way I forget her inertia and consider her the best little housewife in Christendom... Let me close the passage by thanking God for her, and expressing the ever stronger determination to guard and cherish her for ever.

Married life clearly suited him, and it was also a successful time for his mathematical researches.

10th February 1856: ...I have succeeded in establishing a very general and very interesting theorem with respect to the surfaces which equally attract a given point. I hope before midsummer to have a very pretty investigation ready for publication...

But it was all too good to last. Shortly after their wedding, Anna began to show signs of advancing tuberculosis. The symptoms became increasingly worse, and Hirst eventually resigned his job at Queenwood to devote himself to her. From 1856–1857 they travelled in the South of France, vainly searching for a cure. While there, he wrote two papers arising out of his earlier work at Göttingen with Gauss and Weber, and these were published in the *Philosophical Magazine*.

At the same time, his mathematical reading continued to be extensive and intelligent. Even if a work was badly written, he would persevere with it because the subject itself mattered to him. William Rowan Hamilton's work on quaternions, Carl Jacobi's *Elliptical functions* (in Latin), and Sartorius von Waltershausen's *Life of Gauss* were among the works on which he commented, often critically:

14th September 1856: For the last week I have been studying Spottiswoode on Determinants in Crelle's Journal. It is obscurely written and badly printed, and hence very laborious to understand; but as I am determined to master the subject, I shall spare no pains...

18th January 1857: ...I have purchased too an admirable work of Euler's, namely his Letters to a German Princess on subjects connected principally with Physics. The most unscientific person could understand them, they are written with wonderful clearness. I wonder a good translation of them has never been used as reading lessons in our schools. His subjects are not so elementary; it is the lucid style that deceives one into the belief that the subject is simple. Therein consists an infallible sign of an able writer...

Eventually they settled in Paris, where he made the acquaintance of a curious old fellow...

21st June 1857: He is at the same time door-keeper, boot and shoe maker, and mathematician!!!! Like most self-educated men, he is extremely opinionated and almost a monomaniac. Nevertheless he is an original and altogether a remarkable shoemaker. He takes great delight in giving me problems to solve, and is disappointed when I solve them correctly. At present he pronounces my solution of the following problem to be incorrect: "A man borrows 300 francs, for which he is to pay interest at the rate of 5 per cent per ann. If he pays 20 francs a month instead of the interest which is really due, how soon will he have repaid the sum borrowed?"...

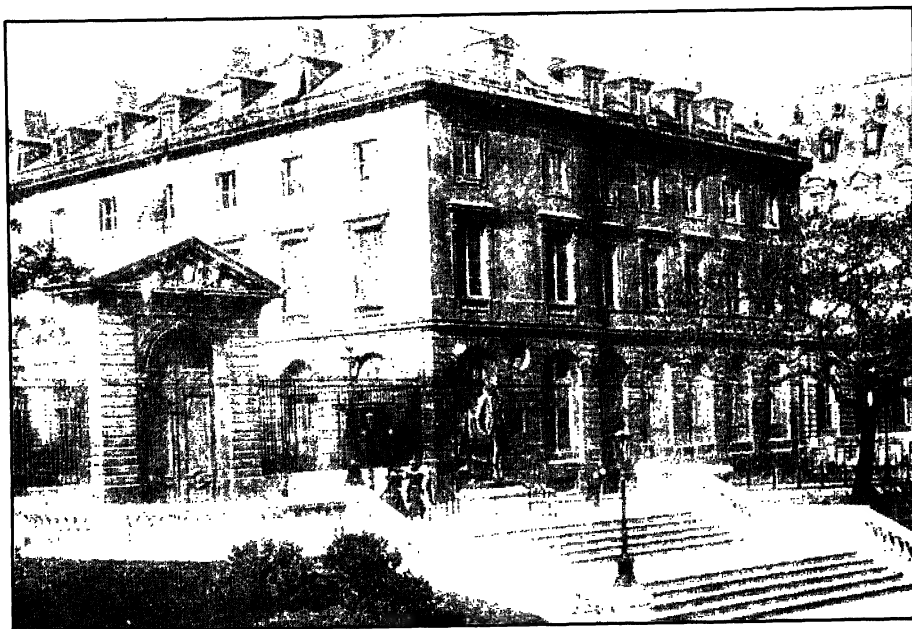
In July the inevitable happened. Anna died, leaving Hirst devastated:

2nd July 1857: Poor Anna suffers no more, she is at peace for ever. Formerly she read my journal and I had always to write accordingly, to leave all my anxieties and fears unexpressed. Now she will read no more...

John Tyndall was on his way to Switzerland to study the structure and movement of glaciers, when he received intelligence of the calamity. He took Hirst with him to Switzerland, and the two became even closer friends. In August 1857 they were joined by their friend Thomas Huxley and made one of the earliest ascents of Mont Blanc.

Hirst never fully got over the tragedy of Anna's death, and paid regular visits to her Paris grave for the rest of his life. Deciding that the time had come to devote himself entirely to research, and perhaps wishing to remain near Anna, he settled himself in Paris.

18th October 1857: My health has continued on the whole good, and I have worked very steadily all the week. Still my progress does not satisfy me. When I consider that I have been nearly two months engaged on a small geometrical research which is a little out of my direct line I feel inclined to lose patience. But I must not. I have commenced a subject and I will give it some kind of finish before I pass to another...



Collège de France

As he became more involved in his researches, he resumed his practice of paying visits to mathematicians. The foremost French mathematicians at this time were Joseph Liouville and Joseph Bertrand at the Collège de France, Michel Chasles at the Sorbonne, and the retired Louis Poincaré.

18th November 1857: On Saturday last I paid M. Liouville a visit. It is long since I first entertained the idea of this visit... He is a pleasant, chatty little man with whom I soon felt at perfect ease. The only blemish I observed in him was an occasional unmeaning giggle. We talked of Dirichlet, of Steiner, of Poincaré, of Cayley and of Sylvester, in the chattiest, frankest manner. His remarks on all these men were shrewd and just. I coincided entirely. And I must confess I heard with some satisfaction his remarks on Cayley's productions. He acknowledged their ability but he protested against their wilful obscurity. He considers Cayley and Sylvester to be in some measure the disciples of Cauchy in this respect... To be precise and clear is equivalent in their eyes to being tedious. Rather than march over their difficulties and through their conquered territory with a firm, steady step, they leap and turn somersaults. It is possible that by so doing *they* are able to take a rapid and sufficient view of their subject, but others decidedly see better with their head upwards...

I went to hear Chasles' first lecture on Geometry, and was far from satisfied with it. Perhaps he was in bad humour—certainly he did not enter with his whole might into his subject. He hesitated and bungled much, and altogether his lecture formed a sad contrast to his books which are remarkably clearly written. But even his books are not to be compared to Steiner's in grasp of his subject...



Joseph Liouville (1809–1882)



Michel Chasles (1793–1880)

Much of Hirst's time was spent in translating important mathematical works into English. One such work was an important memoir on the percussion of bodies by Louis Poinsoot. This gave him the opportunity to visit Poinsoot at his house, where he was met by a footman and conducted to an elegant salon to meet the old man.

20th December 1857: ... He shook me kindly by the hand, bid me be seated, and took his seat near me. He is now between 60 and 70 years old, with silver silken hair neatly arranged on a fine intelligent head. He is tall and thin, but although he now stoops with age and feebleness one can see that one time his figure was more than ordinarily graceful. He was loosely but neatly dressed in a large ample robe de chambre. His features are finely moulded—indeed everything about the man betokens good blood. His eyes are now dim and dull with age, and recede far behind two prominent eyebrows. He talks incessantly and well. I did not misunderstand a word, although he spoke always in a low tone, and now and then his voice dropped as if from weariness, but he never wandered from his point...

Poinsoot was delighted to discuss his works with Hirst, and was clear and interesting in his explanation of them. He seemed touched to hear of his influence on the young Hirst, remarking "We cast our seed upon the waters knowing not where it may fall, but it is nevertheless pleasant after long years of labour to find that these

seeds have taken root.” He presented him with copies of all his works, which pleased the recipient greatly. Hirst obviously read them, for he was soon to write ...

10th January 1858: Without exception Poinso^t’s is the neatest and most lucid mathematical treatise I know. *I find it difficult to put down the book* just as in my younger days I found it difficult to put aside an interesting novel. Poinso^t is one of the few mathematicians who dislike to leave to calculus the task and the merit of arriving at results. With most of us calculation is more than an instrument in our hands, it is a servant in our service to which servant we appoint a task and are but too prone to accept the result he brings to us without enquiring how it has been achieved—Poinso^t on the contrary works *with* this servant, watches his every act and directs the same. The consequence is the result is thoroughly his own ... Every thing he touches he strives to exhaust, he is not satisfied with a simple preception of a truth but he regards it from all sides laboriously and perseveringly until he has found out the path which will lead himself and others most directly and easily to the goal. For young mathematicians I should deem him an admirable instructor.

In January 1858 he received copies of a memoir he had written for Liouville’s Journal, and ‘saw with some little pleasure my name amongst the list of contributors on the cover’, names such as Cayley, Gauss, Jacobi and Dirichlet. But this was not the only exciting event of that month ...

17th January 1858: On the evening of this same day the Emperor of the French [Napoleon III] narrowly escaped assassination at the entrance of the Grand Opera. As usual a crowd was assembled in the Rue Lepeletier to see the arrival of the Emperor and Empress. As their carriage drew up three loud detonations were successively heard, three infernal machines (grenades) exploded under or near his carriage killing and wounding more than a hundred of the spectators, smashing his carriage and slaying one of the horses ...

His mathematical interests now took a new direction, as his work on equally attracting surfaces continually caused him to deviate into geometry.

31st January 1858: ... Having found that two surfaces inscribed in the same cone attract the vertex of the latter equally, provided that radii vectors having the same direction are inversely proportional in length I am led to study what I call *inverse figures* generally. I call two figures inverse with respect to a point O chosen as the centre of inversion, when to every point A of the one corresponds a point A’ of the other so that A and A’ are on a line through O and the rectangle AO. A’O is constant ...

14th February 1858: ... My method of inverse transformation is leading me to a class of curves of the fourth degree which possess properties precisely analogous to but more general than conics. To every theorem in conics concerning points, lines and circles corresponds another with reference to these higher conics concerning points and circles. Conics are as it were turned inside out, their infinitely distant points becoming all concentrated in one point in the plane which I call the point of inversion.

Although he was pursuing his researches in mathematics, he maintained a strong interest in the sciences. He was particularly fascinated by the election to the membership of the Mechanical Section of the Academy of Sciences.

7th March 1858: ... Foucault is a candidate. I noted last night that Chasles will not and Bertrand will support him. His not being a mathematician will in all probability be fatal. Chasles designates his gyroscope and researches on the pendulum as *happy*, but neither indicative of genius or promising in results ...



Joseph Bertrand (1822–1900)

...With respect to Bertrand I am still in doubt whether his harsh, forbidding, arrogant exterior is a true index of his character or merely a cloak to a better nature. To me it is extremely disgusting, the air he assumes. His manner to me appears to repel you by the announcement "what you are telling me may interest you, but as to *me* I knew it all before and much more—in fact with respect to mathematics I am decidedly *blasé*, I may be said to have utterly exhausted that elementary science."

By April, his investigation on equally attracting surfaces was drawing to a close. Although his work had proceeded well, he was unsure of its interest or quality.

25th April 1858: ... It is strange with what different feelings I regard at different times the results of my researches. Sometimes they appear to me of tolerable interest and value, at other times merely curious and common-place. Whatever they may be I hope soon to throw them aside to the indifferent public and occupy myself with others.

6th June 1858: ... I have succeeded in integrating some partial differential equations that have caused me much trouble ... I felt convinced that simple results ought to have been obtained and in fact I found after a while that a mistake where a' was merely put in place of a had caused all the mischief. The thought of the three lost days was as nothing in comparison to the pleasure of seeing complication vanish and former results more than corroborated ...

Ever since his Marburg days his health had caused him problems, which he frequently described in his diaries. In particular, toothache was a recurring problem ...

13th June 1858: I have undergone the very unpleasant operation of burning the nerve. It has changed the nature of the tooth-ache, but not cured it. One night John Martin put me a leech on my gum and it bled profusely for nearly 24 hours ...

Despite such problems, his work progressed well, and by the end of July he had finished his memoir on equally attracting surfaces for the *Philosophical Magazine*.

In August, he left Paris to spend almost a year in Italy—an exciting time to visit, as Italy was in the midst of Civil War.

26th April 1859: ... Austria has declared war to-day against Piedmont. On Saturday last an Austrian Aide de Camp crossed the Ticino to *invite* Sardinia to lay down her arms and disband her volunteers giving her three days to consider her reply. This news appears to be authentic. French troops are quickly moving towards the frontier, it is said they are in Genoa to-day. At any rate a fearful struggle has commenced and God knows how it will end. Its effects will be stamped upon the Century for ever...



Francesco Brioschi (1824–1897)



Luigi Cremona (1830–1903)

Most of his time was spent in Rome working with the mathematician priest Barnaba Tortolini and writing articles for Tortolini's journal. He also met mathematicians in Naples and Sicily. In June 1859 he visited the battlefields near Milan, and witnessed the aftermath of the bloody battle of Solferini.

A few days later, he received a visit from the algebraist Francesco Brioschi.

23rd June 1859: '... he is beyond doubt the ablest mathematician of Italy. He is a rather tall slightly built man with an intelligent earnest face, dark hair and beard and high good forehead, eyes of dark brown in a clear field somewhat sunk but exceedingly intelligent and penetrating... Last Autumn he visited France and Germany and made the acquaintance of the ablest mathematicians of Europe... He deems Cayley about the 1st mathematician of Europe, Hermite the first in France and Kronecker perhaps in Germany. He differed slightly as to the merits of Liouville and some others but agreed perfectly as to Bertrand, Chasles, Steiner &c... In short, of all the mathematicians I have met in Italy he produced upon me the best impression.

Even more important for the future was his first meeting with the geometer Luigi Cremona.

30th June 1859: ... He is a young man, a pupil of Brioschi's, married and has a family. He is short and has a bullet shaped bald head. Our conversation was first of all political and then mathematical; it never flagged and we parted good friends.

After two years abroad, Hirst decided that it was time to return home. After a brief visit to see friends in Marburg and visit Anna's grave in Paris, he set sail for England. The next few years in London were to be the most successful of his career, and form the topic of the next article.

ACKNOWLEDGMENTS. A typed version of the Thomas Hirst diaries is held at the Royal Institution in London, and quotations from the diaries appear here by courtesy of the Royal Institution. The diaries have been edited by W. H. Brock and R. M. MacLeod, and were published in microfiche by Mansell, London, in 1980.

*Open University
Milton Keynes MK7 6AA
England*

ON THE CHINESE ORIGIN OF THE SYMBOL FOR ZERO.

By PROFESSOR FLORIAN CAJORI.

I have just received a letter from Mr. Y. Mikami, of Tokyo, Japan, containing information which (if confirmed by more extended research) is of great interest and importance. The letter is dated December 15, 1902. From it I quote the following:

"I have found very important relations between the mathematics of India and of China. Arabian numerals seem to be of Chinese origin. The abacus, used by the Chinese from time immemorial, probably afforded the principle of position. In China the use of the symbol 0 for zero seems to have been very old. I desire to study the history of the Chinese mathematics from this point of view, if only I can secure sufficient materials, which is, however, very difficult. Chinese works are not [difficult] to understand for us Japanese, because we use the same letters."

Until recently the symbol for zero and the principle of local value in our notation of numbers were supposed to be of Hindu origin. A few years ago our attention was called to the early work of the Japanese, and now the priority appears to be passing to the Chinese.

COLORADO COLLEGE, COLORADO
SPRINGS, *January 3, 1903.*

10(1903), 35

Thoughts on *Innumeracy*: Mathematics Versus the World?

Peter L. Renz

(A reply by John Allen Paulos follows.)

To some, mathematical calculations are soothing and reassuring. The ability to calculate gives them a sense of power. Speaking of an instance in school when his calculation was right and his teacher was wrong, John Allen Paulos wrote:

I remember thinking of mathematics as a kind of omnipotent protector. You could prove things to people and they would have to believe you whether they liked you or not.

(*Innumeracy*, page 73)

Yet his teacher did not believe Paulos's calculation and he didn't acknowledge that Paulos was correct even after seeing Paulos confirmed by figures in the *Milwaukee Journal*.

Calculation has its limits in conquering disbelief, and it has others. As basis for practical decisions or for science, calculation is limited by the accuracy of the data and the correctness of the assumptions on which it is based. Lord Kelvin calculated the age for the Earth based on the rate at which this planet cooled after its formation. He arrived at 20 million years, with 40 million years as a maximum. His calculations were correct; his assumptions were wrong. He did not know of the warming of the Earth's interior by radioactive decay. The current best estimate for the age of the Earth (again a calculation, this one based on radioactive dating) is 4.7 billion years—100 to 200 times the age that Kelvin estimated.

The relentless and immutable nature of calculation, and of mathematics in general, is an affront to some. Among the offended are the circle squarers, the angle trisectors, and the like. These people are John Allen Paulos's innumerates. Their weaknesses lead to diverse problems:

One rarely discussed consequence of innumeracy is its link with belief in pseudoscience.

(*Innumeracy*, page 4)

In addition to astrology, innumerates are considerably more likely than others to believe in visitors from outer space.

(*Innumeracy*, page 59)

... healthy skepticism ... a state of mind generally incompatible with innumeracy.

(*Innumeracy*, page 62)

Paulos gives no quantitative evidence for these commonplace assertions. Paulos

attributes innumeracy to character faults:

Some people personalize events excessively, resisting an external perspective, and since numbers and an impersonal view of the world are intimately related, this resistance contributes to an almost willful innumeracy.

(*Innumeracy*, page 80)

But numeracy helps lift us out of the mire of personal concerns.

If you . . . see happy people holding hands, eating ice cream cones, etc., it's easy to begin to think that other people are happier, more loving, and more productive than you are, and so to become unnecessarily despondent . . . It's beneficial to wonder occasionally what percentage of people you encounter suffer from this or that disease or inadequacy.

(*Innumeracy*, page 81)

There is a hostile and patronizing tone here and an evident lack of sympathy for the innumerate (pity or scorn, yes; sympathy, no). These set my teeth on edge. There is an arrogance and disregard for the difficulties of others and the difficulties of applying mathematics to real problems that reflects poorly on our subject. Consistent with this, *Innumeracy* is flawed by a cavalier disregard for accuracy. Yet despite these faults, this book is a best-seller. Why?

The answer is that we already see innumeracy, however defined, as a general problem (probably in ourselves and certainly in others). Here is a book that confirms a common perception, suggests a ready cure, and does all this with amusing banter and fun number facts. Let me tempt you with this sample:

. . .take a deep breath. Assume Shakespeare's account is accurate and Julius Caesar gasped 'You too, Brutus' before breathing his last. What are the chances that you just inhaled a molecule which Caesar exhaled in his dying breath? The surprising answer is that, with probability better than 99 percent, you did just inhale such a molecule.

(*Innumeracy*, page 24)

Fascinating, and for those who don't believe him Paulos gives the reader a quick calculation to prove his point. Did I believe it? No, and here is why.

Paulos states that the number of molecules in the atmosphere is about 10^{44} . Where did this number come from? I had no idea, and Paulos gives no clues, but by digging around in *The Handbook of Physics and Chemistry* I found figures for the mass of the atmosphere and the molecular constitution of the atmosphere that made his number a reasonable estimate. Next, Paulos states that a breath is $\frac{1}{30}$ th of a liter and contains 2.2×10^{22} molecules. As we shall see, this is wrong on two counts. First, a gram molecular weight (mole) of any gas at standard temperature and pressure fills 22.4 liters and contains 6×10^{23} molecules. The number 6×10^{23} is Avogadro's number, the number of molecules in a mole of any compound. A quick calculation shows that Paulos should have gotten

$$\frac{1}{30} \times \frac{1}{22.4} \times 6 \times 10^{23} = 8.9 \times 10^{20}$$

molecules per breath instead of 2.2×10^{22} . But let's follow his calculation as he made it, using his number of molecules per breath. Suppose all the molecules in Caesar's last breath are uniformly mixed up in today's atmosphere. (Is this reasonable?) To get a handle on this, let's call the molecules from Caesar's last breath "lucky" and all other molecules "unlucky." The probability of a random

molecule's being lucky is just

$$\frac{\text{Number of lucky molecules}}{\text{Number of molecules in atmosphere}} = \frac{2.2 \times 10^{22}}{10^{44}} = 2.2 \times 10^{-22}$$

The probability of a random molecule's being unlucky is

$$\begin{aligned} \frac{\text{Number of unlucky molecules}}{\text{Number of molecules in atmosphere}} &= \frac{10^{44} - 2.2 \times 10^{22}}{10^{44}} \\ &= 1 - \frac{2.2 \times 10^{22}}{10^{44}} \\ &= (1 - 2.2 \times 10^{-22}) = Q. \end{aligned}$$

We call this number Q .

The probability of two random molecules being unlucky is effectively $Q \times Q$. (The second draw is not independent of the first because this is sampling without replacement. Calculation shows that the adjustment for dependence leaves the first twenty or so significant figures unaffected and can be neglected. Paulos makes no mention of this, although dependence can be important and the observation that it can be neglected here is a nice exercise in approximation.) The probability of your whole lungful of molecules (all 2.2×10^{22} of them according to Paulos) consisting only of unlucky molecules is then just

$$P = Q^{2.2 \times 10^{22}} = (1 - 2.2 \times 10^{-22})^{2.2 \times 10^{22}}.$$

Paulos tells his reader that this product is less than 0.01. True, but how would an even moderately sophisticated reader calculate $(1 - 2.2 \times 10^{-22})^{2.2 \times 10^{22}}$? You can't use your pocket calculator because $1 - 2.2 \times 10^{-22}$ figured on a calculator is 1 and the exponent is out of range. Repeated multiplication is out of the question; it would take too long. You must use natural logs or the definition of e . Either approach uses calculus and yields

$$P = e^{-4.84} \approx 0.0079.$$

This is Paulos's probability of a whole lungful of unlucky molecules. So his probability of at least one lucky molecule in a random lungful is $1 - P = 1 - 0.0079 = 0.992$ or better than 99%.

What are my complaints? First, no innumerate (and relatively few numerates) could fill in the steps. Second, Paulos's numbers are wrong. If you use his $\frac{1}{30}$ th of a liter per breath, the calculation gives the probability of a random breath's not containing a molecule of Caesar's last breath as

$$P' = \left(1 - \frac{8.9 \times 10^{20}}{10^{44}}\right)^{8.9 \times 10^{20}} = 0.992.$$

So the probability of getting a lungful of unlucky molecules is 0.992. (By coincidence, this number matches one in Paulos's calculation, but it gives the complementary probability.) Continuing, the probability of getting at least one lucky molecule in a lungful is $1 - 0.992 = 0.008$, or less than 1%—contrary to what Paulos writes.

What does this tell us? First, you get wrong answers from bad numbers. Second, when simple operations like addition, subtraction, multiplication, and raising to a power are taken to extremes, special techniques must be used. Third, it is not easy to dig up good values for the numbers needed in many calculations.

This calculation is mentioned without details in J. E. Littlewood's *A Mathematician's Miscellany*, and Littlewood credits James Jeans. I tracked this to Jeans's *An Introduction to the Kinetic Theory of Gases*, Cambridge University Press, 1942. With a breath of 0.4 liter, 10^{22} molecules, this is also the number of such breaths in the atmosphere, which Jeans puts at 10^{44} molecules. With proper mixing, each breath could contain a molecule of Caesar's last breath. No fanfare. Jeans's numbers are good and his calculation is immediate. Paulos's calculation is tricky and his volume for a breath too small. The volume is close to 1/2 liter (more for a deep breath). Paulos did not check the volume of a breath either by experiment or in references. I looked at *Human Respiration* by Olof Lippold, W. H. Freeman and Company, San Francisco, 1968 and I experimented as well.

The hypothesis of random mixing of the molecules of Caesar's last breath in the atmosphere is dubious. There is no evidence that Paulos checked this. There are several problems concerning this mixing. Molecules of air dissociate and can recombine forming other molecules or react to become part of the biosphere, hydrosphere, or even end up in sediment. Looking into this requires a bit of research. Nitrogen is the main constituent of air (80% of it). The amount of nitrogen in sediment is more than that in the atmosphere. However, interchange between atmosphere and sediment is quite slow. One must check on this. My source was Delwiche's article "The Nitrogen Cycle," *Scientific American*, September 1970. These numbers are rough, but they suggest that it is safe to assume almost all the nitrogen molecules in Caesar's last breath are still in the atmosphere, but it does to speak to the uniform mixing of those molecules in the atmosphere.

We can work out Paulos's calculation with an average breath of 1/2 liter, assuming total random mixing of the original molecules of Caesar's last breath in the atmosphere, and that there is no loss of those molecules. The probability of a random breath's not containing any "lucky" molecules is

$$\left(1 - \frac{1.3 \times 10^{22}}{10^{44}}\right)^{1.3 \times 10^{22}} = e^{-1.3 \times 1.3} = 0.16.$$

So the probability of getting at least one lucky molecule in an average lungful is

$$1 - 0.16 = 0.84.$$

By increasing the estimated size of a breath of air you can pump this probability up to Paulos's 99%.

There is a final question here: What is the purpose of such a calculation? Is the object simply to amaze the reader, or is it to instruct, or is it intended to lead to some course of action? What do we learn from Paulos here? Jeans and Littlewood, speaking to those who could work out the technicalities, had clear points in mind, but Paulos's purpose is unclear.

Reviewing *Innumeracy* in *The Washington Post*, Eleanor Wilson Orr, a mathematics and science teacher for 35 years and an author writing on issues in mathematics education, said, "... for the innumerate who wants to take this book seriously and read it carefully, the book is intimidating. ... I learned a lot from this book but I spent five full days reading it with a pencil in my hand. I fiddled with the numbers, I drew diagrams, I daydreamed and tried to explain to myself what Paulos doesn't explain. I trusted that I would understand it if I kept at it. Innumerates either quit or think it's enough to get the general idea, and so remain innumerate." She noted none of Paulos's errors. Judith Axler Turner, who wrote an article on Paulos and his book for the *Chronicle of Higher Education*, com-

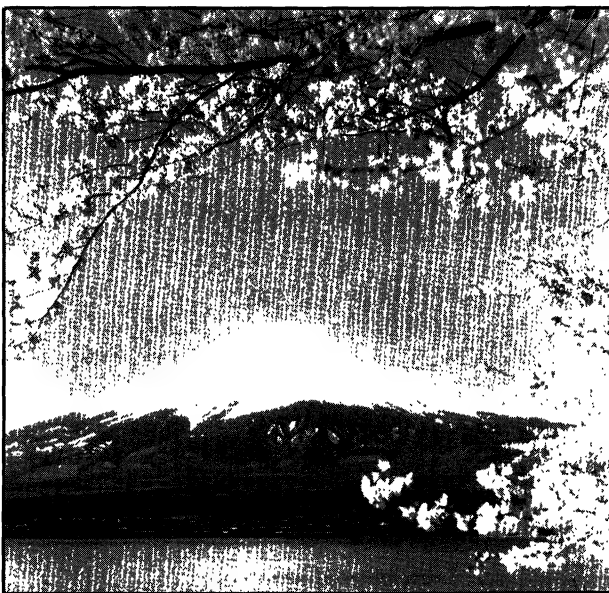
mented that Paulos scoffed at Orr's difficulties, but I do not scoff. The last of Orr's sentences quoted in on the mark. You cannot read Paulos's book seriously without giving some attention to the details and that attention will require serious work. Not only will it require serious work but that work will reveal that there is less in Paulos's book than meets the eye of the casual reader.

Here is another numerical problem Paulos poses and answers. It is equally amusing but seems more practical.

One last earthly calculation that a scientific consultant from M.I.T. uses to weed out prospective employees during job interviews: How long, he asks, would it take dump trucks to cart away an isolated mountain, say Japan's Mount Fuji, to ground level? Assume trucks come every fifteen minutes, twenty-four hours a day, are instantaneously filled with mountain dirt and rock, and leave without getting in each other's way. The answer is a little surprising and will be given later.

(*Innumeracy*, page 12)

The answer, without explanation, appears in a sentence on page 15 where Paulos estimates it would take 5,000 to 10,000 years to truck away Mount Fuji. This is a surprisingly short time for such a job. It is also wrong. The only fact that Paulos mentions about Mount Fuji is its height, 12,000 feet, so it is clear that he figured the mountain was some sort of cone. The volume of a cone is a third of its base area multiplied by its height, a fact easily derived and known to Archimedes. Evidently, Paulos must have also used the area of the base of Mount Fuji in his calculation. Did he look this up? Did he consult maps? No, as an exchange of letters revealed, he dreamed it up. He assumed Mount Fuji was a cone as wide at its base as it was high. Volcanos are simply not shaped this way, and one might expect Paulos, who spent a year at the University of Washington within easy view of Mount Rainier, to know something about the shape of a volcano. Leaving



The gentle slopes of Mount Fuji are shown here. Photograph courtesy of the Japan National Tourist Organization.

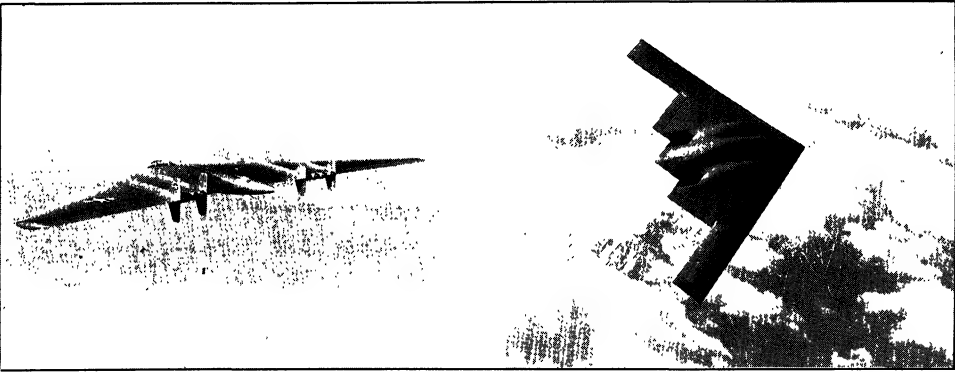
that aside, you might expect him, as an author, to look at an atlas. I did. The map is revealing. It shows that Fuji is roughly conical and has a radius of about 12 kilometers at its 1000 meter contour. Its height is 3776 meters above sea level. Below 1000 meters it broadens out considerably. We might construe the problem of trucking away Mount Fuji as that of taking enough of it away so that what was left would blend into the countryside. From the map it looks as if taking the top 2776 meters off the peak would do the job. The volume of that part of the

mountain is

$$\frac{1}{3} \times \pi (12,000 \text{ m})^2 \times 2776 \text{ m} = 4.19 \times 10^{11} \text{ m}^3.$$

Calling a local importer of heavy Japanese trucks, I found that the largest standard model that they imported could carry 18.5 cubic yards. Round up to 20 cubic meters per load, and divide by the product—cubic meters per load times loads per hour times hours per day, etc.—and you will find that it would take about 600,000 years to truck away the top 2776 meters of Mount Fuji. To cart away a cone the same shape and the height of Mount Fuji measured from sea level (3776 meters) would take over 1.5 million years at this rate. Paulos’s estimate of 5,000 to 10,000 years is off by orders of magnitude. Would his mythical M.I.T.-based recruiter have hired him for some practical job? I would hope not, but given the errors committed in real-world engineering, perhaps so. Note that even with these considerations, this is a highly idealized problem. It is clear that no such project could ever be carried out.

Here is an example of real-life erroneous calculation with a potentially large impact. These calculations were made by William R. Sears and Irving L. Ashkenas in a secret assessment of promising aeronautical technologies that they prepared in 1945. Sears and Ashkenas built a mathematical model to show how the range of an aircraft varied as one redistributed the volume between the wing and the fuselage. Sears and Ashkenas were working at the time for the Northrop Corporation, a firm then building various experimental “flying wing” aircraft. They differentiated their formula for range as a function of the percentage of volume in the wings and found only two possible extrema: one of these was when all the volume was in the wing and the other when a much smaller fraction of the volume was in the wing. Sears and Ashkenas wrote, “It can be ascertained that the form [all volume in the wing] gives maximum range, while the latter gives a minimum.” Hence, flying wings have the maximal range.



The Northrop YB-49 Flying Wing, left, and its sleek delta-form descendant, the B-2 Stealth bomber, right. Both pictures courtesy of the Northrop Corporation.

Joseph Foa, who headed a group studying possible designs for an unmanned jet aircraft at Cornell Aeronautical Laboratory (CAL) had reached the contrary conclusion—that a flying wing configuration would not give maximum range. After Sears came to head CAL, Foa had a chance to examine the Sears-Ashkenas report and discovered that the critical point associated with the flying wing configuration

was a minimum, not a maximum. It gave the minimum range according to their model, not the maximum range, as Sears and Ashkenas claimed.

This came to light in *Science* (Volume 244, pp. 650–651, 12 May 1989) in connection with controversy about the B-2 stealth bomber, also a flying wing. In the 1940s Foa kept his silence on the condition that Sears and Ashkenas publish some sort of correction to their earlier analysis. That correction took the form of the 1948 paper “Range performance of turbojet airplanes” by Ashkenas in the *Journal of the Aeronautical Sciences*. Here Ashkenas made a much more complex mathematical model, which had the property that the flying wing configuration gave optimal range for certain choices of the basic parameters. To this day Foa remains unconvinced, asserting that the Ashkenas optimum flying wing would be impractically thick.

The B-2 project had a multi-billion dollar budget. The initial error of Sears and Ashkenas, mistaking a minimum for a maximum, is a classic for students in freshman calculus. But even after careful consideration by competent aeronautical engineers, it is not clear whether the flying wing is the best or the worst way to go if one wants a long-range plane. The answer you get from a mathematical model seems to depend on what answer you want to get.

Our quantitative understanding of the world is not simply based on assumptions; it is based on observation. Generally there is a lot of hard work needed to get good numbers. When it comes to projecting cancers that may or may not be caused by the breakdown products of minute quantities of Alar in apples, the work is hard, the numbers are soft, and the theoretical apparatus is quite involved. The meaning of such calculations is more controversial, important, and uncertain than for the calculations I have discussed above. Paulos gives scant attention to any of this.

Paulos is quick to point out the problems of others.

A recent study by Drs. Kronlund and Phillips of the University of Washington showed that doctors’ assessments of the risks of various operations, procedures, and medications (even in their own specialties) were way off the mark, often by several orders of magnitude.

(*Innumeracy*, page 8.)

Paulos continues in the cited paragraph, heaping scorn on doctors, “I once had a conversation with a doctor who, within approximately twenty minutes, stated that a certain procedure he was contemplating (a) had a one-chance-in-a-million risk associated with it; (b) was 99% safe; and (c) usually went quite well. Given the fact that so many doctors seem to believe that there must be at least eleven people in the waiting room if they’re to avoid being idle, I am not surprised at this new evidence of their innumeracy.” Let’s think this through. If (a) is true, then (b) follows, because it is a weaker condition. Furthermore, it is reasonable (c) might also be true. A doctor might say that such a procedure “usually went well,” although this is a qualitative judgment having to do with ease of the procedure and lack of difficulties for the doctor. Now, I expect the doctor in question did not have much detailed statistical information, because such information is difficult and costly to gather. But does what Paulos present show the doctor to be innumerate? Not at all, and the line about waiting rooms is a cheap shot meant to please readers who have cooled their heels in a doctor’s office.

The importance and the difficulty of gathering good data on medical matters are greater than one might think. Let me use an example from *Innumeracy*. Paulos begins a calculation on page 21 by stating that the probability of heterosexual

transmission of AIDS from an infected to an uninfected person is 0.002 per act of intercourse. This probability is called the infectivity. He says this is an average of figures from several studies—but he cites no sources. This makes me wonder, because it is difficult to see how one could get good figures on the transmissibility of this disease. Only a Dr. Mengele operating without restraint could plan and execute experiments on AIDS infectivity. The problems include the facts that AIDS is sexually transmitted and 100% fatal. It is extremely difficult to get accurate information about sexual behavior. These matters are private and sensitive. People regularly lie about sexual matters, and Congress regularly kills publicly-funded studies of sexual behavior.

I asked experts, including Eric Lander who organized the National Academy of Sciences session on AIDS, about reliable information on the heterosexual transmissibility of AIDS and came up with little. The best source I found was the issue of *Los Alamos Science*, Number 18, 1989, devoted to AIDS. The lead article, “AIDS and a Risk-Based Model,” by Colgate, Stanley, Hyman, Qualls, and Layne, gives estimates for the infectivity ranging from 0.0014 to 0.004. These authors cite others whose estimates of the infectivity run from 0.003 to 0.1. The methods discussed are difficult and use epidemiological data and complex assumptions. There is a factor of about 100 between the largest and smallest of these estimates of infectivity, and Paulos’s 0.002 falls within the range, on the low side. It would be prudent not to put too much faith in any particular number for the infectivity.

Let’s see what use Paulos makes of this probability. He says we may assume these probabilities of transmission to be independent. He then notes that $(1 - 0.002)^{346} \approx 0.5$. So that a year’s worth of nightly unprotected intercourse with an infected partner leaves you with a better than 50% of being uninfected. Next, he asserts that if a condom is used the infectivity drops to 2×10^{-4} . Now you can enjoy ten years of nightly intercourse with the victim (assuming the victim lives this long, Paulos adds parenthetically) before your probability of getting AIDS rises to 0.5. Finally, Paulos states that the probability of contracting AIDS from a single act of unprotected intercourse with someone belonging to no known risk group is 2×10^{-7} and with a condom this probability drops to 2×10^{-8} . He writes that you will more likely die in a car accident returning from such a tryst than catch AIDS during the act. All this suggests that one need not worry all that much about AIDS. Now these calculations are correct, though the assumptions underlying them are dubious, and the suggestion that AIDS is not very worrisome is dead wrong.

AIDS transmission is extremely variable. It appears that an individual can be so infectious as to infect virtually everyone with whom he has unprotected sexual intercourse. The evidence for this comes from an Australian sperm donor. His frozen sperm sample was split into ten doses, eight of which were used, resulting in four infected women. A Poisson model is suggested for assaying infectivity by dilution methods in “The Kinetics of HIV Infectivity,” by Layne, Dembo, and Spouge in the cited issue of *Los Alamos Science*. Let N be the number of individuals treated with the diluted infectious agent (here, $N = 8$). Let d be the dilution that infects 50% of the individuals treated (here $d = 0.1$), and I be the infectivity of the undiluted semen. Then in this instance

$$\text{Number infected} = 4 = 0.5N = Ne^{-dI} = Ne^{-0.1I}.$$

The exponential factor comes from the Poisson probability,

$$p_k = e^{-dI} (dI)^k / k!$$

with $k = 0$. Using this we can estimate the probability of infection from a single act of intercourse with this donor at the time of donation

$$1 - \text{Probability of no infection} = 1 - e^{-I} = 1 - \frac{1}{2^{10}} \approx 0.999$$

where the probability of no infection is simply obtained from the Poisson p_0 with $d = 1$ and I evaluated from $0.5 = e^{-dI}$.

Not much data is available. But to illustrate the variability of transmission of AIDS, I quote another example. Sperm from an infected New York donor was used to inseminate 90 women, none of whom contracted AIDS. Even given selectivity in reporting, it is unlikely that the cited Australian and New York examples are samples drawn from the same population.

What of Paulos's comment about the risk of being killed in an automobile accident versus the risk of getting AIDS? It is a common and generally uncalculated guess that the risks of doing X are less than the risk of driving to the place where you do X. From Paulos's figures, it might be the other way around in this case. The average US passenger death rate as given in *The World Almanac and Book of Facts*: 1992 is 1.12×10^{-8} deaths per mile. Compare to Paulos's estimate of a risk of 2×10^{-8} for contracting AIDS from protected intercourse with someone having no known risk factors. The risk that dominates will depend, at least, on how far away the tryst is. More importantly, the death rate per mile depends strongly on the age, sex, and driving history of the driver and on such things as sobriety, roads and road conditions, and on whether a seat belt is used. Your risk per mile could be quite a bit larger or smaller than the average which I gave. As we begin to bring in these considerations, we move from a general statistical treatment toward special cases and special pleading. More data would be needed to establish the risks for these new classes. This leads away from easy calculation and toward more specialized cases. This sort of thinking is a poor guide for public policy, but we live or die as special cases, not on the average.

This simply hints at what is wrong with the lax, breezy treatment Paulos gives. When applied to so serious a matter as AIDS, it is shocking. Yes, innumeracy is a problem, but *Innumeracy* is more a part of the problem than of the solution. This need not have been the case. Were the book less negative toward the innumerate and more carefully done, it could have made a wonderful contribution. My copy has quite a few favorable comments in the margins along with many notes on errors of the sort I mentioned here.

We should hold ourselves, our students, and others to higher standards. We want public appeal, clarity, and *truth*. This will not be easy to get, but why settle for less?

101 Colchester Street
Brookline, MA 02146

John Allen Paulos replies:

A distant relative of mine recently returned from a three-month stay in Florida. When I asked him how it was, he launched into a detailed disquisition on the mechanical minutiae of his new appliance for extracting juice from oranges and grapefruits. He took offense at my attempt to summarize his comments or to use

Reflections on Rippling Water

Michel Mendes France

1. INTRODUCTION. On a summer evening standing beside a large lake that extends to the horizon, we observe the moon's reflection on the rippled surface of the water. When the moon is low, but nonetheless completely above the horizon, the reflection may still appear as a long uninterrupted yellow column which stretches from some point on the lake to the horizon. Its length can be considered as infinite. Later, when the moon rises higher up in the sky the reflection changes aspect and becomes a shorter beam, in the shape of a narrow oval. It is closer to us and no longer extends to the horizon. Its length is now finite.

Stars may appear in this evening sky. If a gentle breeze is blowing, each one of these stars will appear to be reflected an odd number of times in a given direction on the surface of the lake.

These evocative images raise interesting mathematical questions. At what angle does the moon's reflection change from an infinite image to a finite one? Is it possible to see exactly two reflections of the same star? The object of this paper is to answer these questions. Our analysis only requires simple trigonometry.

2. THE THEORY. Suppose an observer at height H_1 sees a reflected object on the wavelet M at a distance x across the water.

Let $\alpha = \alpha(x)$ be the angle measured in radians between the normal MN to the wave with the vertical V . Let α_0 be the maximal value of $|\alpha(x)|$ and define $\varphi(x)$ by $\alpha(x) = \alpha_0 \varphi(x)$ so that $|\varphi(x)| \leq 1$. Let i be the angle of reflection (Figure 1 and 2).

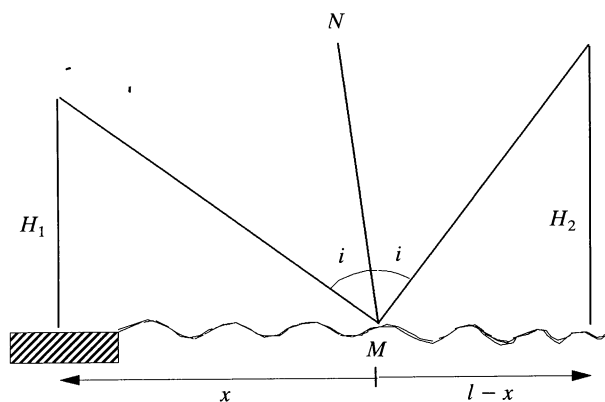


Figure 1

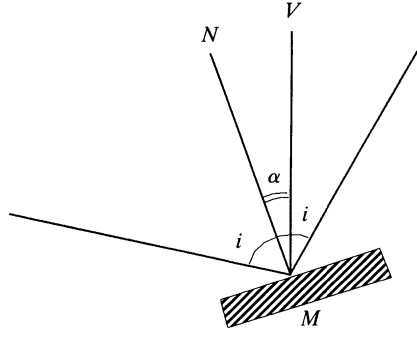


Figure 2

Trivial trigonometry shows that

$$\begin{cases} \tan(i + \alpha) = \frac{x}{H_1} \\ \tan(i - \alpha) = \frac{l - x}{H_2} \end{cases}.$$

Hence

$$\begin{aligned} \tan 2\alpha &= \tan[(\alpha + i) - (i - \alpha)] \\ &= \frac{\tan(\alpha + i) - \tan(i - \alpha)}{1 + \tan(\alpha + i)\tan(i - \alpha)} \\ &= \frac{\frac{x}{H_1} + \frac{x - l}{H_2}}{1 + \frac{x(l - x)}{H_1 H_2}} = \frac{(H_1 + H_2)x - lH_1}{H_1 H_2 + lx - x^2}. \end{aligned}$$

We now assume that α_0 is small. Then

$$2\alpha \approx \frac{(H_1 + H_2)x - lH_1}{H_1 H_2 + lx - x^2}.$$

Finally,

$$\varphi(x) \approx \frac{(H_1 + H_2)x - lH_1}{2\alpha_0(H_1 H_2 + lx - x^2)}. \quad (1)$$

Before exploiting the relationship given by (1), let us analyze the corresponding equation

$$\varphi(x) = \frac{x(H_1 + H_2) - lH_1}{2\alpha_0(H_1 H_2 + lx - x^2)}. \quad (2)$$

Note that each side of (2) has a physical interpretation. The function φ describes the shape of the waves while the right-hand side represents the distance at which a reflection occurs. So, for a given shape φ , the solutions of (2) are the approximate distances at which a reflection occurs. In particular, the number of solutions is the number of reflected images we see.

Now let us analyze the equation. We start by looking at the simplest case when there are no waves at all: $\varphi \equiv 0$. Then

$$x = \frac{lH_1}{H_1 + H_2}.$$

Thus there is only one reflection so that the observer sees a perfect image of the object. If, in particular $H_1 = H_2$, then $x = l/2$ and the light ray is reflected at the midpoint between the observer and the object. This is of course well known.

Let us now discuss the general case where φ stays small. We solve the equation (2) graphically.

Let β be the curve

$$\beta: x \mapsto \frac{x(H_1 + H_2) - lH_1}{2\alpha_0(H_1H_2 + lx - x^2)}.$$

In the interval $(0, l)$, β is continuous and increasing. Furthermore

$$\beta(0) = -\frac{l}{2\alpha_0H_2} \quad \text{and} \quad \beta(l) = \frac{l}{2\alpha_0H_1}.$$

We assume that both H_1 and H_2 are strictly less than $l/2\alpha_0$ (l is large and α_0 is small).

The curve $x \mapsto \varphi(x)$ oscillates in the horizontal strip $y = -1, y = +1$. Supposing φ is continuous in the interval $[0, l]$, both curves intersect either at an odd number of points or infinitely often. Thus, whatever the shape of the waves may be, one should see either an odd number of reflections, or infinitely many. (This last case may indeed occur if, for example, φ has a singularity of the type $(x - a)\sin(x - a)^{-1}$ in the neighbourhood of some $a \in (0, l)$).

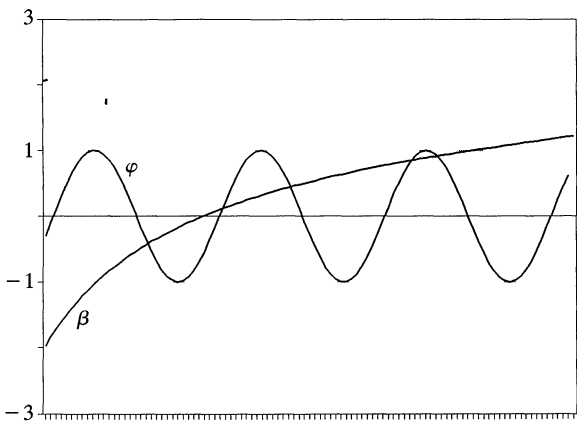


Figure 3

3. LIMITING CASES. Let us study the solutions of equation (2) when $l = +\infty$ (reflection of the moon or the sun...).

We denote by w the angle at which the infinitely far away object is seen. Then H_1/l is negligible and $H_2/l = \tan w$. Thus rewriting the right hand side of equation (1) in the form

$$\frac{x(H_1 l^{-1} + H_2 l^{-1}) - H_1}{2\alpha_0 \left(-\frac{x^2}{l} + x + \frac{H_1 H_2}{l} \right)}$$

we have

$$\varphi(x) \approx \frac{x \tan w - H_1}{2\alpha_0(x + H_1 \tan w)} = \gamma(x).$$

As before we solve the equation $\varphi(x) = \gamma(x)$ graphically and we suppose that φ oscillates a great many times, say

$$\varphi(x) = \sin \lambda x$$

where λ is large. We solve equation (3) for $x \in (0, l)$

$$\sin \lambda x = \gamma(x). \tag{3}$$

Since $\gamma(x)$ is monotonically increasing we know that the smallest solution x_S of (3) occurs when this function is -1 and the largest solution x_L occurs when the function is $+1$.

When $\tan w < 2\alpha_0$, we see from Figure 4 that the two curves intersect infinitely many times and the smallest solution is approximately

$$x_S \approx \max \left\{ 0, H_1 \frac{1 - 2\alpha_0 \tan w}{\tan w + 2\alpha_0} \right\}.$$

In this case, the reflection on the water extends from x_S to the horizon. When $\tan w > 2\alpha_0$, the solution is shown on Figure 5.

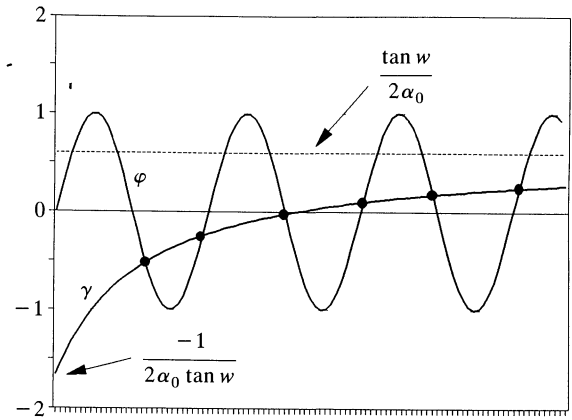


Figure 4

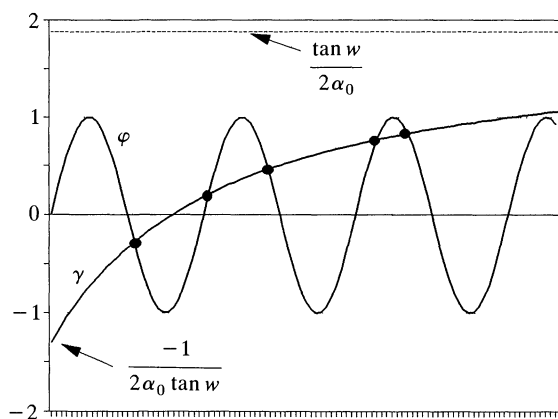


Figure 5

In this case there is only a finite odd number of reflections and the reflections lie between x_S and x_L

$$x_L \approx \frac{H_1}{\tan w - 2\alpha_0} (1 + 2\alpha_0 \tan w).$$

It follows that the critical w_c at which the reflection ceases to be infinite is therefore

$$w_c = \tan^{-1}(2\alpha_0).$$

As α_0 is assumed to be small, we have

$$w_c \approx 2\alpha_0.$$

Finally, if one is given the shape of the waves as a Cartesian equation

$$y = \psi(x),$$

then

$$\alpha(x) = \tan^{-1}\psi'(x),$$

provided ψ is differentiable. As $|\alpha(x)|$ is small, this entails $\alpha(x) \approx \psi'(x)$ so that the critical angle is

$$w_c \approx 2 \max_x |\psi'(x)|.$$

4. AN APPLICATION. The amplitudes of real waves on the ocean, far away from the coast (say 100 yards or more) are difficult to measure: they move rapidly and their size may be small, especially if we are discussing wavelets or even ripples. On the other hand, w_c can be quite easily measured at sunset: observe at what angle w_c the reflection starts to touch the horizon. If we assume that during that time of waiting, the waves keep approximately the same shape, say

$$\varphi(x) = \alpha_0 \sin \lambda x \cos \lambda ct$$

where c is the velocity of the wave and t is time, then the knowledge of w_c and of the frequency λ gives us the amplitude A of the waves

$$A = \frac{w_c}{2\lambda}.$$

Our analysis is also valid for studying the microscopic structure of a glossy surface. The macroscopic observation of a reflecting luminous point provides information on the fine structure of the surface. Determining w_c measures the product $A\lambda$.

It was only after completing this work that I discovered M. Minnaert's delightful book [1] on the "Nature of Light & Colour in the open air." It discusses related topics and I highly recommend it (see in particular pp. 23–26). I wish to thank the referee and Jacques Harthong for helping me to improve the exposition and the graphs.

Addendum. Many authors have studied the reflection on rippling water. I would like to single out M. V. Berry's beautiful article "Disruption of images: the caustic-touching theorem," *J. Opt. Soc. Am. A*, 4, 1987, pp. 561–569.

REFERENCE

1. M. Minnaert, *The Nature of Light & Colour in the Open Air*, Dover Publ., Inc., 1954.

*Department of Mathematics
Université Bordeaux I
F-33405 Talence Cedex
France*

PICTURE PUZZLE
(from the collection of Paul Halmos)



Are they related?
(see page 809.)

The Principal Axis Theorem over Arbitrary Fields

David Mornhinweg, Daniel B. Shapiro, and K. G. Valente

The Principal Axis Theorem, included in most undergraduate texts in Linear Algebra though often without proof, states that every symmetric matrix over the field of real numbers is orthogonally similar to a diagonal matrix. In [1], S. Friedberg, focusing attention on the underlying field, gave an elementary argument to show that there are symmetric matrices over \mathbf{Z}_p (p a prime) which are not orthogonally similar to a diagonal matrix. This paper concludes with a problem: “Classify exactly those fields for which the Principal Axis Theorem is true.” As solutions to this classification problem can be found in the literature (see [2] and [10] for example) and frequent reconsiderations of this topic indicate an interest to a wide audience of mathematicians, the purpose of this paper is to give a simplified overview of this beautiful result. As we proceed, we keep an eye toward the accessibility of the argument. In fact, with the exception of two technical results, this development can be incorporated in any undergraduate-level course in Linear Algebra that deals with arbitrary fields. For example, one can show quite easily that it is necessary for the field to have characteristic equal to zero in order to insure that symmetric matrices are diagonalizable. While it is a rather straightforward matter to establish a large class of fields which allows for the orthogonal diagonalization of symmetric matrices, one of the aforementioned technical results is crucial in the final step of the classification of such fields.

A field F is said to have the *Principal Axis Property* if every symmetric matrix over F is orthogonally similar to a diagonal matrix over F . That is, for every symmetric matrix M over F , there exists an orthogonal matrix P over F (that is, $P^{-1} = P^t$) such that $P^{-1}MP$ is diagonal.

A study of the 2×2 case provides some important information. We write $\text{char}(F)$ for the characteristic of F .

Lemma 1. *Suppose every symmetric 2×2 matrix over F is diagonalizable over F . Then*

- (i) $\sqrt{-1} \notin F$,
- (ii) every sum of squares in F is a square in F , and
- (iii) $\text{char}(F) = 0$.

Proof: Just suppose there exists $i \in F$ with $i^2 = -1$. Then the matrix

$$\begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$$

has characteristic polynomial x^2 . If this matrix were diagonalizable, it would have to be the zero matrix. This contradiction establishes (i). As an immediate consequence we have $\text{char}(F) \neq 2$, for if $\text{char}(F) = 2$, then $i = 1 = -1$ in F .

To prove (ii) it suffices to show that if $a, b \in F$ then $a^2 + b^2$ is a perfect square in F . To see this we consider the matrix

$$M = \begin{bmatrix} a & b/2 \\ b/2 & 0 \end{bmatrix}$$

which has eigenvalues $\frac{1}{2}(a \pm \sqrt{a^2 + b^2})$. Since M is diagonalizable over F , we know that these eigenvalues lie in F , so that $\sqrt{a^2 + b^2} \in F$. Property (iii) now follows from (i) and (ii), for if $\text{char}(F) = p > 0$ then $-1 = (p-1)$ is a sum of squares in F . \square

Every field satisfying the conditions of Lemma 1 must possess an “ordering”. To explain why this is so, we outline some of the properties of ordered fields. These ideas were introduced by E. Artin and O. Schreier in the 1920’s and have since appeared in many algebra texts. The basic idea is that an order relation on F , which respects the field operations of F , is determined by the “cone” P of non-negative elements. This P is taken as the fundamental object.

Definition. An *ordering* on a field F is a subset $P \subseteq F$ satisfying (i) $P + P \subseteq P$; (ii) $P \cdot P \subseteq P$; (iii) $P \cap -P = \{0\}$; (iv) $P \cup -P = F$.

Here $-P := \{-a : a \in P\}$. Given an ordering P , we define an order relation \leq on F as follows: $a \leq b$ if and only if $b - a \in P$. The reader is invited to derive the familiar properties of “less-than-or-equal” from the given axioms. For example, $a^2 \geq 0$ for every element a . This follows from (ii) if $a \in P$. Otherwise, $a \notin P$ and (iv) implies that $-a \in P$, so that $a^2 = (-a)^2 \in P$. From (i) it follows that every sum of squares in F must lie in P . This proves that if F admits an ordering, then F must be “formally real” in the following sense.

Definition. A field F is *formally real* if -1 is not expressible as a sum of squares in F .

From our remarks above, we see that the complex field \mathbb{C} has no ordering. Also if $\text{char}(F) > 0$, then F has no ordering. However some fields admit several orderings. For instance if $\sigma : F \rightarrow \mathbb{R}$ is a homomorphism into the field of real numbers, then $P = \sigma^{-1}([0, \infty))$ is an ordering on F . Distinct embeddings of F into \mathbb{R} yield distinct orderings of F . For example $\mathbb{Q}(\sqrt{2})$ possesses two orderings.

Our first technical result completes the connection between ordered and formally real fields.

Theorem 1. *A field F admits an ordering if and only if F is formally real.*

This famous theorem was first proved by Artin in the 1920’s as part of his solution to Hilbert’s 17th problem. The construction of an ordering on a formally real field invokes the Axiom of Choice and appears in a number of texts, including [7], [8] and [9].

The second property appearing in Lemma 1 has also been given a name.

Definition. A field F is *pythagorean* if every sum of squares in F is a square in F .

The fields of real and complex numbers are pythagorean, while the field of rationals is not. Using this new terminology, Lemma 1 can be restated as follows: if 2×2 matrices over F can be diagonalized over F , then F must be formally real and pythagorean. Therefore, in our search for fields which satisfy the Principal Axis Property, we may restrict our attention to formally real pythagorean fields. We now point out that such fields allow for the Gram-Schmidt orthogonalization process on F^n .

Lemma 2. *Let F be a formally real pythagorean field and let \geq be any order relation on F . If $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ in F^n , define*

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + \cdots + u_n v_n.$$

Then this map $\langle \cdot, \cdot \rangle: F^n \times F^n \rightarrow F$ is an inner product, and for every $\mathbf{v} \in F^n$ there exists a unique element $\|\mathbf{v}\| \in F$ with $\|\mathbf{v}\| \geq 0$ and $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$.

Proof: For $\mathbf{v} \in F^n$ we see that $\langle \mathbf{v}, \mathbf{v} \rangle = v_1^2 + \cdots + v_n^2 \geq 0$ in F . Moreover if $\mathbf{v} \neq \mathbf{0}$ then some $v_i \neq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle \neq \mathbf{0}$. It follows that $\langle \cdot, \cdot \rangle$ is an inner product. Since F is pythagorean we know that $\langle \mathbf{v}, \mathbf{v} \rangle$ is a square in F . Every square in F has a unique non-negative square root in F , and the lemma follows. \square

We are now in a position to rephrase the question found in [5].

Theorem 2. *Let F be a formally real pythagorean field. The following are equivalent:*

- (i) *F has the Principal Axis Property,*
- (ii) *Every symmetric matrix over F is diagonalizable over F , and*
- (iii) *Every symmetric matrix over F has an eigenvalue in F .*

Proof: Clearly (i) \Rightarrow (ii) \Rightarrow (iii). To prove (iii) \Rightarrow (i) we let M be an $n \times n$ symmetric matrix over F and proceed by induction on n . The result is clear when $n = 1$ so we may assume $n > 1$. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard orthonormal basis of F^n . By (iii) the matrix M has an eigenvalue $k \in F$, and there exists a corresponding eigenvector $\mathbf{w} \in F^n$. Complete this vector to a basis of F^n , and apply the Gram-Schmidt process to obtain an orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$, where $\mathbf{u}_1 = \mathbf{w}/\|\mathbf{w}\|$. Let P be the matrix with columns equal to these vectors \mathbf{u}_i and let $S = P^{-1}MP$. Then P is an orthogonal matrix since the columns form an orthonormal basis. Therefore S is also symmetric. The first column of S is $S\mathbf{e}_1 = P^{-1}MP\mathbf{e}_1 = P^{-1}M\mathbf{u}_1 = k \cdot P^{-1}\mathbf{u}_1 = k\mathbf{e}_1$. The first row of S is then determined by the symmetry and we see that

$$S = \begin{pmatrix} k & 0 & 0 & \cdots & 0 & 0 \\ 0 & & & & & \\ 0 & & & & & \\ \vdots & & & S_0 & & \\ 0 & & & & & \\ 0 & & & & & \end{pmatrix}$$

where S_0 is a symmetric $(n-1) \times (n-1)$ matrix. By induction, there exists an $(n-1) \times (n-1)$ matrix R_0 such that $R_0^{-1} = R_0'$ and $R_0^{-1}TR_0 = D$ where D is a

diagonal matrix. Now, set

$$R = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & & & & & \\ 0 & & & & & \\ \vdots & & & R_0 & & \\ 0 & & & & & \\ 0 & & & & & \end{pmatrix}$$

and consider the matrix PR . We see that $(PR)^{-1} = (PR)^t$, and

$$(PR)^{-1}M(PR) = \begin{pmatrix} k & 0 & 0 & \cdots & 0 & 0 \\ 0 & & & & & \\ 0 & & & & & \\ \vdots & & & D & & \\ 0 & & & & & \\ 0 & & & & & \end{pmatrix}. \quad \square$$

We note that when working over the field of reals or real algebraic numbers one can appeal to the standard arguments involving the Fundamental Theorem of Algebra to conclude by Theorem 2 that all symmetric matrices can be diagonalized. In particular, both of these fields are real closed. For our purposes, a field F is *real closed* if it is pythagorean, formally real, and $F(\sqrt{-1})$ is algebraically closed. (There are many other equivalent definitions for real closed.) This class of fields was also studied by Artin and Schreier and further information regarding these fields can be found in the aforementioned texts.

This theorem also implies that an intersection of fields satisfying the Principal Axis Property again has that property, although some care must be taken to ensure that the intersection makes sense. To guarantee that the field operations are compatible, we assume that all the fields in question are subfields of some larger field.

Corollary.

- (i) *Any real closed field satisfies the Principal Axis Property.*
- (ii) *Let Ω be a field with $\{F_\alpha\}$ a collection of subfields. If each field F_α satisfies the Principal Axis Property, then their intersection also satisfies the Principal Axis Property.*

Proof: Using the definition of real closed given above, the standard argument establishing the existence of a real eigenvalue for a real symmetric matrix can be adapted to prove (i). For a more complete development of diagonalization over real closed fields, one can also see [9].

To prove (ii) let $F = \bigcap_\alpha F_\alpha$ be the intersection. By Lemma 1 and the subsequent definitions, we know that each F_α is formally real and pythagorean, and therefore so is F . If M is a symmetric matrix over F , then Theorem 2 implies that all of the eigenvalues of M lie in F_α . Since this is true for every α , we see that the eigenvalues lie in F . The claim follows by another application of Theorem 2. \square

With this corollary we see that the intersection of any collection of real closed fields (that are subfields of a common field) satisfies the Principal Axis Property.

In fact, these are only fields with this property. To see this we are in need of a second technical result due to F. Krakowski [6]. As before, we must assume all the fields under consideration lie inside some larger field. To this end, let F be a field with Ω a fixed algebraically closed extension of F . Set

$$R(F) = \bigcap \{K|F \subseteq K \subseteq \Omega \text{ and } K \text{ is real closed}\}.$$

Note, using Theorem 1, if F is not formally real, then $R(F)$ is trivial. On the other hand, if F is formally real, then $R(F)$ is a formally real pythagorean extension of F . Further information regarding the construction of $R(F)$ can be found in [3], [7] and [10].

Theorem 3. *Let F be a formally real pythagorean field. For any $a \in R(F)$, there exists a symmetric matrix over F having a as an eigenvalue.*

Proof: (Sketch) Let V denote the field $F(a)$ with $B : V \times V \rightarrow F$ the trace form. That is,

$$B(x, y) = \text{tr}(xy)$$

where tr is the trace mapping from V to F . Let $T : V \rightarrow V$ be the F -linear map defined by $T(b) = ab$. Since $B(T(x), y) = B(x, T(y))$, T is self-adjoint with respect to the symmetric bilinear form B . By our choice of a , B is positive definite with respect to every possible ordering of F . In other words, in any diagonal representation of B , every diagonal entry must be a sum of squares in F and therefore a square as F is pythagorean. With this, one can choose a basis for V so that the matrix for B with respect to this basis is the identity matrix. Letting S represent the matrix of T relative to this basis, the self-adjoint behavior of T implies that S is symmetric. By construction, a is an eigenvalue of T and therefore an eigenvalue of S . \square

The proof of this theorem shows that if $a \in R(F)$ has degree n over F then a is an eigenvalue of some symmetric $n \times n$ matrix. For a more general field F and $a \in R(F)$ it is interesting to ask what size of symmetric matrix is required to have a as an eigenvalue. For example, for the field \mathbb{Q} of rational numbers, $R(\mathbb{Q})$ is the set of real algebraic numbers. In [1], E. Bender showed that every real algebraic number of degree n is an eigenvalue of some symmetric $(n + 1) \times (n + 1)$ matrix over \mathbb{Q} . The analogous question for algebraic integers as eigenvalues of symmetric integer matrices has been considered by D. Estes [4].

With this result we can now give a complete characterization of the fields for which the Principal Axis Property holds.

Theorem 4. *A field F satisfies the Principal Axis Property if and only if F is an intersection of real closed fields.*

Proof: The “if” part is established by the Corollary to Theorem 2. To continue, let $a \in R(F)$. Choosing a symmetric matrix over F having a as an eigenvalue, we see that $a \in F$ by hypothesis. Thus $F = R(F)$ and the characterization is complete. \square

ACKNOWLEDGMENTS. The authors wish to thank E. Becker and A. Wadsworth for their help in revising this paper. Their interest and suggestions are greatly appreciated.

1. E. A. Bender, The dimensions of symmetric matrices with a given minimum polynomial, *Linear Alg. Appl.* 3 (1970) 115–123.
2. A. Charnow and E. Charnow, Fields for which the principal axis theorem is valid, *Math. Mag.* 59 (1986) 222–225.
3. T. Craven, Intersections of real closed fields, *Can. J. Math.* 32 (1980) 431–440.
4. D. Estes, Eigenvalues of symmetric integer matrices, (preprint 1991).
5. S. H. Freidberg, Extending the principal axis theorem to fields other than \mathbb{R} , *Amer. Math. Monthly* 97 (1990) 147–149.
6. F. Krakowski, Eigenwerte und Minimalpolynome symmetrischer Matrizen in kommutativen Körpern, *Comment. Math. Helv.* 32 (1958) 224–240.
7. T. Y. Lam, *The Algebraic Theory of Quadratic Forms*, W. A. Benjamin Publ. Co., Reading, MA, 1973.
8. W. Scharlau, *Quadratic and Hermitian Forms*, Springer Verlag, Berlin-Heidelberg-New York-Tokyo, 1985.
9. B. L. van der Waerden, *Algebra*, vol. 2, Frederick Ungar Publ. Co., New York, 1970.
10. W. Waterhouse, Self-adjoint operators and formally real fields, *Duke Math. J.* 43 (1976) 237–243.

Department
of Mathematics
Colgate University
Hamilton, NY 13346

Department
of Mathematics
Ohio State University
Columbus, OH 43210

Department
of Mathematics
Colgate University
Hamilton, NY 13346

THE CHAUVENET PRIZE.

The committee on the award of the first Chauvenet Prize for excellence in mathematical exposition, Professors W. C. GRAUSTEIN, ANNA PELL WHEELER, and A. B. VAN VLECK, chairman, recommended that the award be made to Professor G. A. BLISS of the University of Chicago for his paper on “Algebraic functions and their divisors,” published in the *Annals of Mathematics*, volume 26, Numbers 1 and 2, September and December 1924. The Trustees voted to approve this choice and to thank the members of the committee for their arduous but very valuable efforts. The award was announced at the business meeting and the prize of one hundred dollars, furnished by a member of the Association, was presented to Professor Bliss following the meetings.

33(1926), 177

The Fifty-Third William Lowell Putnam Mathematical Competition

Leonard F. Klosinski
Gerald L. Alexanderson
Loren C. Larson

The following results of the fifty-third William Lowell Putnam Mathematical Competition, held on December 5, 1992, have been determined in accordance with the governing regulations. This annual contest is supported by the William Lowell Putnam Prize Fund for the Promotion of Scholarship, left by Mrs. Putnam in memory of her husband, and is held under the auspices of the Mathematical Association of America.

The first prize, \$7,500, was awarded to the Department of Mathematics of Harvard University. The members of the winning team were: Jordan S. Ellenberg, Samuel A. Kutin, and Royce Y. Peng; each was awarded a prize of \$500.

The second prize, \$5,000, was awarded to the Department of Mathematics of the University of Toronto. The members of the winning team were: J. P. Grossman, Jeff T. Higham, and Hugh R. Thomas; each was awarded a prize of \$400.

The third prize, \$3,000, was awarded to the Department of Mathematics of the University of Waterloo. The members of the winning team were Dorian Birsan, Daniel R. L. Brown, and Ian A. Goldberg; each was awarded a prize of \$300.

The fourth prize, \$2,000, was awarded to the Department of Mathematics at Princeton University. The members of the winning team were Joshua B. Fischman, Adam M. Logan, and Joel E. Rosenberg; each was awarded a prize of \$200.

The fifth prize, \$1,000, was awarded to the Department of Mathematics at Cornell University. The members of the winning team were Jon M. Kleinberg, Mark Krosky, and Demetrio A. Muñoz; each was awarded a prize of \$100.

The five highest ranking individual contestants, in alphabetical order, were Jordan S. Ellenberg, Harvard University; Samuel A. Kutin, Harvard University; Adam M. Logan, Princeton University; Serban M. Nacu, Harvard University; and Jeffrey M. Vanderkam, Duke University. Each of these was designated a Putnam Fellow by the Mathematical Association of America and awarded a prize of \$1,000 by the Putnam Prize Fund.

The next six highest ranking contestants, in alphabetical order, were David B. Carlton, Harvard University; Ian A. Goldberg, University of Waterloo; Kiran S. Kedlaya, Harvard University; Royce Y. Peng, Harvard University; Hugh R. Thomas, University of Toronto; and Tong Zhang, Cornell University; each was awarded a prize of \$500.

The next four highest ranking individuals, in alphabetical order, were Ze-Yu Chen, Princeton University; Jonathan T. Higa, Princeton University; Svetlozar E. Nestorov, Stanford University; and Samuel K. Vandervelde, Swarthmore College; each was awarded a prize of \$250.

The next nine highest ranking individuals, in alphabetical order, were Daniel R. L. Brown, University of Waterloo; Jeff T. Higham, University of Toronto; F. Dean Hildebrandt, Harvard University; Julie B. Kerr, Washington State University; Andrew H. Kresch, Yale University; William R. Mann, Princeton University; Dana Pascovici, Dartmouth College; Michail G. Sunitsky, Princeton University; and Douglas J. Zare, New College of the University of South Florida; each was awarded a prize of \$100.

The following teams, named in alphabetical order, received honorable mention: Dartmouth College, with team members Radu Bacioiu, Rolf H. Nelson, and Dana Pascovici; Duke University, with team members Craig B. Gentry, Alexander J. Hartemink, and Jeffrey M. Vanderkam; Massachusetts Institute of Technology, with team members Thomas C. Chou, Henry L. Cohn, and Michael J. Lawler; University of British Columbia, with team members Malik H. Kalfane, David L. Savitt, and Mark A. Van Raamsdonk; and Yale University, with team members Thomas Feng, Andrew H. Kresch, and Zhaohui Zhang.

Honorable mention was achieved by the following thirty-one individuals named in alphabetical order: James McCleery Berger, Brown University; Sergey Brin, University of Maryland, College Park; Thomas C. Chou, Massachusetts Institute of Technology; Henry L. Cohn, Massachusetts Institute of Technology; Brian D. Ewald, University of Michigan, Ann Arbor; Joshua B. Fischman, Princeton University; J. P. Grossman, University of Toronto; Steven S. Gubser, Princeton University; William M. Hesse, University of Connecticut; Adam Kalai, Harvard University; Timothy P. Kokesh, Harvey Mudd College; Botond Kőszegi, Harvard University; Peter R. Kramer, Princeton University; Mark Krosky, Cornell University; Tal N. Kubo, Harvard University; Sergey V. Levin, Harvard University; Samuel J. Maltby, University of Calgary; Demetrio A. Muñoz, Cornell University; Akira Negi, University of North Carolina, Chapel Hill; Seth Padowitz, Brown University; Andrew Przeworski, Massachusetts Institute of Technology; Philip T. Reiss, University of Manitoba; James P. Sarvis, Massachusetts Institute of Technology; Kannan Soundararajan, University of Michigan, Ann Arbor; Michael G. Szydlo, Boston University; Joe Y. Tien, University of California, Irvine; Mark A. Van Raamsdonk, University of British Columbia; Jeffrey D. Wall, Princeton University; Kelly Lynne Wieand, University of Wisconsin, Madison; Erick B. Wong, Simon Fraser University; and Zhaohui Zhang, Yale University.

The other individuals who achieved ranks among the top 98, in alphabetical order of their schools, were: Brigham Young University, John Wesley Robertson; University of British Columbia, David L. Savitt; Brown University, Andrew Brecher; California Institute of Technology, Steven C. Anderson; University of California, Berkeley, Daniel C. Isaksen; University of Colorado, Boulder, Steve T. Soulé; Cornell University, Jon M. Kleinberg; Dartmouth College, Radu Bacioiu; Duke University, Alexander J. Hartemink; Harvard University, Manjul Bhargava, Joseph I. Chuang, Michael L. Hutchings, Dimitri Kountourogiannis, Paul Li, Matteo J. Paris, Chris Ternoey; Harvey Mudd College, Jon H. Leonard; University of Maine, Orono, YuQun Chen; Massachusetts Institute of Technology, Jerome S. Khohayting, Tichomir G. Teney, William W. Tucker; Memorial University of Newfoundland, Robert P. Gallant; Michigan State University, Thomas P. Hayes; University of Minnesota, Minneapolis, Matthew P. Kelly; Université de Montréal, Marc-André

Lafortune; New York University, Mikhail Kogan; Ohio State University, Frank J. Swenton; University of Pennsylvania, Frosti Petursson; Princeton University, Tibor Beke, Mark W. Lucianovic; Purdue University, Pok-Yin Yu; Rice University, Donald A. Barkauskas; Rose Hulman Institute of Technology, Jonathan E. Atkins; Stanford University, Daniel P. Cory, Garrett R. Vargas; Texas A & M University, Zheng-Zheng Li; University of Waterloo, Dorian Birsan, Kevin K. Cheung, Jie J. Lou; Wellesley College, Yihao L. Zhang; West Virginia Wesleyan College, Emanuel V. Todorov; and Yale University, Matthew Frank.

The Elizabeth Lowell Putnam Prize, named for the wife of William Lowell Putnam and to be “awarded periodically to a woman whose performance on the Competition has been deemed particularly meritorious”, is awarded this year for the first time to Dana Pascovici of Dartmouth College. The winner is awarded a prize of \$500.

There were 2421 individual contestants from 393 colleges and universities in Canada and the United States in the competition of December 5, 1992. Teams were entered by 284 institutions.

The Questions Committee for the fifty-third competition consisted of George E. Andrews (Chair), George T. Gilbert, and Eugene Luks; they composed the problems listed below and were most prominent among those suggesting solutions.

PROBLEMS

Problem A-1.

Prove that $f(n) = 1 - n$ is the only integer-valued function defined on the integers that satisfies the following conditions:

- (i) $f(f(n)) = n$, for all integers n ;
- (ii) $f(f(n + 2) + 2) = n$ for all integers n ;
- (iii) $f(0) = 1$.

Problem A-2.

Define $C(\alpha)$ to be the coefficient of x^{1992} in the power series expansion about $x = 0$ of $(1 + x)^\alpha$. Evaluate

$$\int_0^1 C(-y - 1) \left(\frac{1}{y + 1} + \frac{1}{y + 2} + \frac{1}{y + 3} + \cdots + \frac{1}{y + 1992} \right) dy.$$

Problem A-3.

For a given positive integer m , find all triples (n, x, y) of positive integers, with n relatively prime to m , which satisfy $(x^2 + y^2)^m = (xy)^n$.

Problem A-4.

Let f be an infinitely differentiable real-valued function defined on the real numbers. If

$$f\left(\frac{1}{n}\right) = \frac{n^2}{n^2 + 1}, \quad n = 1, 2, 3, \dots,$$

compute the values of the derivatives $f^{(k)}(0)$, $k = 1, 2, 3, \dots$.

Problem A-5.

For each positive integer n , let

$$a_n = \begin{cases} 0 & \text{if the number of 1's in the binary representation of } n \text{ is even,} \\ 1 & \text{if the number of 1's in the binary representation of } n \text{ is odd.} \end{cases}$$

Show that there do not exist positive integers k and m such that

$$a_{k+j} = a_{k+m+j} = a_{k+2m+j}, \quad \text{for } 0 \leq j \leq m-1.$$

Problem A-6.

Four points are chosen at random on the surface of a sphere. What is the probability that the center of the sphere lies inside the tetrahedron whose vertices are at the four points? (It is understood that each point is independently chosen relative to a uniform distribution on the sphere.)

Problem B-1.

Let S be a set of n distinct real numbers. Let A_S be the set of numbers that occur as averages of two distinct elements of S . For a given $n \geq 2$, what is the smallest possible number of distinct elements in A_S ?

Problem B-2.

For nonnegative integers n and k , define $Q(n, k)$ to be the coefficient of x^k in the expansion of $(1 + x + x^2 + x^3)^n$. Prove that

$$Q(n, k) = \sum_{j=0}^n \binom{n}{j} \binom{n}{k-2j},$$

where $\binom{a}{b}$ is the standard binomial coefficient. (Reminder: For integers a and b with $a \geq 0$, $\binom{a}{b} = a!/(b!(a-b)!)$ for $0 \leq b \leq a$, and $\binom{a}{b} = 0$ otherwise.)

Problem B-3.

For any pair (x, y) of real numbers, a sequence $(a_n(x, y))_{n \geq 0}$ is defined as follows:

$$\begin{aligned} a_0(x, y) &= x, \\ a_{n+1}(x, y) &= \frac{(a_n(x, y))^2 + y^2}{2}, \quad \text{for all } n \geq 0. \end{aligned}$$

Find the area of the region $\{(x, y) \mid (a_n(x, y))_{n \geq 0} \text{ converges}\}$.

Problem B-4.

Let $p(x)$ be a nonzero polynomial of degree less than 1992 having no nonconstant factor in common with $x^3 - x$. Let

$$\frac{d^{1992}}{dx^{1992}} \left(\frac{p(x)}{x^3 - x} \right) = \frac{f(x)}{g(x)}$$

for polynomials $f(x)$ and $g(x)$. Find the smallest possible degree of $f(x)$.

Problem B-5.

Let D_n denote the value of the $(n-1) \times (n-1)$ determinant

$$\begin{vmatrix} 3 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 4 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 5 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 6 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & \cdots & n+1 \end{vmatrix}.$$

Is the set $\{D_n/n!\}_{n \geq 2}$ bounded?

Problem B-6.

Let \mathcal{M} be a set of real $n \times n$ matrices such that

- (i) $I \in \mathcal{M}$, where I is the $n \times n$ identity matrix;
- (ii) if $A \in \mathcal{M}$ and $B \in \mathcal{M}$, then either $AB \in \mathcal{M}$ or $-AB \in \mathcal{M}$, but not both;
- (iii) if $A \in \mathcal{M}$ and $B \in \mathcal{M}$, then either $AB = BA$ or $AB = -BA$;
- (iv) if $A \in \mathcal{M}$ and $A \neq I$, there is at least one $B \in \mathcal{M}$ such that $AB = -BA$.

Prove that \mathcal{M} contains at most n^2 matrices.

SOLUTIONS

In the 12-tuples $(n_{10}, n_9, \dots, n_0, n_{-1})$ following each problem number below, n_i for $10 \geq i \geq 0$ is the number of students among the top 203 contestants achieving i points for the problem and n_{-1} is the number of those not submitting solutions.

A-1 (31, 82, 42, 10, 0, 0, 0, 7, 23, 6, 2, 0)

Solution. If $f(n) = 1 - n$, then $f(f(n)) = f(1 - n) = 1 - (1 - n) = n$, so (i) holds. Similarly, $f(f(n+2)+2) = f((-n-1)+2) = f(1-n) = n$, so (ii) holds. Clearly (iii) holds, and so $f(n) = 1 - n$ satisfies the conditions.

Conversely, suppose f satisfies the three given conditions. From condition (ii), $f(f(f(n+2)+2)) = f(n)$, and applying (i) yields $f(n+2)+2 = f(n)$ or $f(n+2)+2 = -f(n)$.

2) = $f(n) - 2$. An easy induction yields

$$f(n) = \begin{cases} f(0) - n & \text{if } n \text{ is even,} \\ f(1) + 1 - n & \text{if } n \text{ is odd.} \end{cases}$$

If $f(0) = 1$, then $f(1) = 0$ by (i), therefore, $f(n) = 1 - n$.

A-2 (157, 1, 0, 0, 0, 0, 0, 2, 14, 14, 15)

Solution. From the binomial series, we see that

$$\begin{aligned} C(-y-1) &= \frac{(-y-1)(-y-2) \cdots (-y-1992)}{1992!} \\ &= \frac{(y+1)(y+2) \cdots (y+1992)}{1992!}. \end{aligned}$$

Therefore,

$$\begin{aligned} C(-y-1) &\left(\frac{1}{y+1} + \frac{1}{y+2} + \cdots + \frac{1}{y+1992} \right) \\ &= \frac{d}{dy} \left(\frac{(y+1)(y+2) \cdots (y+1992)}{1992!} \right). \end{aligned}$$

Hence the integral in question is

$$\begin{aligned} \int_0^1 \frac{d}{dy} \left(\frac{(y+1)(y+2) \cdots (y+1992)}{1992!} \right) dy &= \frac{(y+1)(y+2) \cdots (y+1992)}{1992!} \Big|_0^1 \\ &= 1993 - 1 = 1992. \end{aligned}$$

A-3 (55, 20, 7, 0, 0, 0, 0, 16, 7, 45, 53)

Solution. There are no solutions if m is odd. If m is even, the only solution is $(n, x, y) = (m+1, 2^{m/2}, 2^{m/2})$.

If (n, x, y) is a solution, then by the arithmetic-mean—geometric-mean inequality, $(xy)^n = (x^2 + y^2)^m \geq (2xy)^m$, so $n > m$. Let p be a prime number. Let a and b be the largest powers of p that divide x and y , respectively. Then the largest power of p dividing $(xy)^n$ is $(a+b)n$. If $a < b$, the largest power of p dividing $(x^2 + y^2)^m$ is $2am$. But this implies that $(a+b)n = 2am$, and this contradicts $n > m$. Similarly, the assumption $a > b$ leads to a contradiction. Therefore $a = b$ for all primes p , and we conclude that $x = y$. Thus, the equation reduces to $(2x^2)^m = x^{2n}$, or equivalently, $x^{2(n-m)} = 2^m$. It follows that x is a positive power of 2, say 2^a . This implies $2(n-m)a = m$, or, $2an = (2a+1)m$. Since $\gcd(m, n) = \gcd(2a, 2a+1) = 1$, we must have $m = 2a$ and $n = 2a+1$. Thus, m is necessarily even and the solution follows as claimed.

A-4 (17, 6, 7, 0, 0, 0, 2, 0, 73, 18, 47, 33)

Solution. We will show that

$$f^{(k)}(0) = \begin{cases} (-1)^{k/2} k! & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

First we note that if $h(x)$ is a differentiable function and x_1, x_2, \dots , is a sequence strictly decreasing to 0 such that $h(x_n) = 0$, then by Rolle's Theorem, there exists a sequence y_1, y_2, \dots , strictly decreasing to 0, such that $h'(y_n) = 0$ ($x_{n+1} < y_n < x_n$).

Now let $g(x) = f(x) - 1/(1+x^2)$. Then $g(1/n) = 0$ for $n = 1, 2, \dots$. Applying the result of the preceding paragraph to g, g', g'', \dots and invoking the continuity of $g^{(k)}$ at 0, we see that $g^{(k)}(0) = 0$ for $k = 0, 1, 2, 3, \dots$. Thus,

$$f^{(k)}(0) = \frac{d^k}{dx^k} \left(\frac{1}{1+x^2} \right) \Big|_{x=0}.$$

The Maclaurin series for $1/(1+x^2)$ is $\sum_{k=0}^{\infty} (-1)^k x^{2k}$, and hence $f^{(k)}(0)$ is equal to the values given above.

A-5 (1, 9, 1, 0, 0, 0, 0, 5, 3, 72, 112)

Solution. Observe that $a_{2n} = a_n$ and $a_{2n+1} = 1 - a_{2n} = 1 - a_n$.

Suppose that there exist k, m as above, and we may assume m is minimal for such.

Suppose first that m is odd. We'll suppose $a_k = a_{k+m} = a_{k+2m} = 0$, as it will be clear that the case $a_k = 1$ can be treated similarly. Since either k or $k+m$ is even, $a_{k+1} = a_{k+m+1} = a_{k+2m+1} = 1$. Again, since either $k+1$ or $k+m+1$ is even, $a_{k+2} = a_{k+m+2} = a_{k+2m+2} = 0$. By this means, we see that the terms $a_k, a_{k+1}, a_{k+2}, \dots, a_{k+m-1}$ alternate between 0 and 1. Then since $m-1$ is even, $a_{k+m-1} = a_{k+2m-1} = a_{k+3m-1} = 0$. But, since either $k+m-1$ or $k+2m-1$ is even, that would imply that $a_{k+m} = a_{k+2m} = 1$, a contradiction.

Thus, m must be even. Extracting the terms with even indices in

$$a_{k+j} = a_{k+m+j} = a_{k+2m+j}, \quad \text{for } 0 \leq j \leq m-1,$$

and using the fact that $a_r = a_{r/2}$ for even r , we get

$$a_{\lfloor k/2 \rfloor + i} = a_{\lfloor k/2 \rfloor + (m/2) + i} = a_{\lfloor k/2 \rfloor + m + i}, \quad \text{for } 0 \leq i \leq (m/2) - 1.$$

(The even numbers $\geq k$ are $2\lfloor k/2 \rfloor, 2\lfloor k/2 \rfloor + 2, \dots$.) This contradicts the minimality of m .

Hence, there are no such k and m .

A-6 (9, 3, 4, 0, 0, 0, 0, 0, 10, 32, 22, 123)

Solution. Recall first that if points A, B, C, D are in general position in 3-space, then a point E lies inside the tetrahedron $ABCD$ if and only if the barycentric coordinates of E with respect to A, B, C, D are positive. That is, if we (uniquely) express

$$\vec{E} = w\vec{A} + x\vec{B} + y\vec{C} + z\vec{D}, \quad \text{with } w + x + y + z = 1,$$

(the arrows indicating consideration of the coordinate triples as vectors), then E is in the interior of $ABCD$ if and only if $w > 0, x > 0, y > 0$, and $z > 0$. Hence, if E is the origin, then E is in the interior of $ABCD$ if and only if there is a solution (w, x, y, z) to

$$\vec{0} = w\vec{A} + x\vec{B} + y\vec{C} + z\vec{D} \tag{1}$$

with w, x, y, z having the same sign. As the solution space to (1) is 1-dimensional, this condition holds for one nonzero solution if and only if it holds for all.

Now assume that the center of the sphere is located at the origin and fix the first chosen point P on the sphere as the north pole, the other three points, P_1, P_2, P_3 , then being random.

We may suppose the choice of each P_i is made in two steps, the first choosing a random diameter $Q_{i_1}Q_{i_2}$ and the second choosing at random between the endpoints Q_{i_1}, Q_{i_2} . Since the $2^3 = 8$ possible selections of endpoints of the three diameters are equally likely, each of the 8 tetrahedra $PQ_{1j_1}Q_{2j_2}Q_{3j_3}$, $j_i = 1$ or 2 , are equally likely. We may further suppose that the vertices of each of these tetrahedra are in general position as the probability of degeneracy is 0. Similarly, we may suppose that the center of the sphere does not lie on any face of the tetrahedra.

Let (w, x, y, z) be a nonzero solution to the equation

$$\vec{0} = w\vec{P} + x\vec{Q}_{11} + y\vec{Q}_{21} + z\vec{Q}_{31}.$$

Then, since $\vec{Q}_{i1} = -\vec{Q}_{i2}$, the eight equations

$$\vec{0} = w\vec{P} + x\vec{Q}_{1j_1} + y\vec{Q}_{2j_2} + z\vec{Q}_{3j_3}$$

have respective solutions

$$(w, x, y, z), (w, x, y, -z), (w, x, -y, z), (w, -x, y, z),$$

$$(w, x, -y, -z), (w, -x, -y, z), (w, -x, y, -z), (w, -x, -y, -z).$$

Hence, exactly one of the eight equations has a solution whose coordinates have the same sign.

It follows that exactly one of these 8 equally likely tetrahedra contains the center. Thus the probability of including the center is $1/8$ for all initial choices of 3 diameters. We conclude that the probability for a random tetrahedron is $1/8$.

B-1 (145, 15, 4, 0, 0, 0, 0, 6, 14, 11, 8)

Solution. The smallest possible number of elements in A_S is $2n - 3$.

Let $x_1 < x_2 < \dots < x_n$ represent the elements of S . Then

$$\begin{aligned} \frac{x_1 + x_2}{2} &< \frac{x_1 + x_3}{2} < \dots < \frac{x_1 + x_n}{2} < \frac{x_2 + x_n}{2} < \frac{x_3 + x_n}{2} \\ &< \dots < \frac{x_{n-1} + x_n}{2} \end{aligned}$$

represent $(n - 1) + (n - 2) = 2n - 3$ distinct elements of A_S , so A_S has at least $2n - 3$ distinct elements.

On the other hand, if we take $S = \{1, 2, \dots, n\}$, the elements of A_S are $\frac{3}{2}, \frac{4}{2}, \frac{5}{2}, \dots, \frac{2n-1}{2}$. There are only $(2n - 1) - 2 = 2n - 3$ such numbers; thus there is a set A_S with at most $2n - 3$ distinct elements. This completes the proof.

B-2 (159, 10, 7, 0, 0, 0, 0, 0, 1, 4, 13, 9)

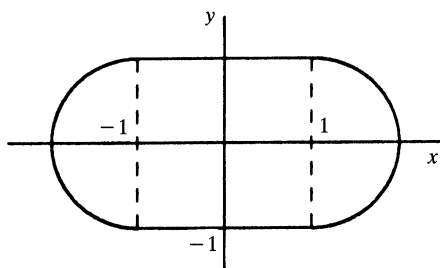
Solution. We have

$$\begin{aligned}
 \sum_{k \geq 0} Q(n, k) x^k &= (1 + x + x^2 + x^3)^n \\
 &= (1 + x^2)^n (1 + x)^n \\
 &= \sum_{j \geq 0} \binom{n}{j} x^{2j} \sum_{i \geq 0} \binom{n}{i} x^i \\
 &= \sum_{j \geq 0} \sum_{i \geq 0} x^{2j+i} \binom{n}{j} \binom{n}{i} \\
 &= \sum_{k \geq 0} x^k \sum_{j \geq 0} \binom{n}{j} \binom{n}{k-2j}.
 \end{aligned}$$

Comparing coefficients of x^k , we derive the desired result.

B-3 (23, 11, 10, 0, 0, 0, 0, 0, 27, 24, 71, 37)

Solution. The area is $4 + \pi$. The region of convergence is



namely, a (closed) square $\{(x, y) \mid -1 \leq x, y \leq 1\}$ of side 2 with (closed) semicircles of radius 1 centered at $(\pm 1, 0)$ described on two opposite sides.

If $\lim_{n \rightarrow \infty} a_n(x, y) = L$, then L must satisfy $L = (L^2 + y^2)/2$; that is, L must be a root of the equation

$$r^2 - 2r + y^2 = 0. \quad (1)$$

In such case, the equation must have real roots, so the discriminant, $4 - 4y^2$, must be nonnegative. Thus, a necessary condition for $(a_n(x, y))$ to converge is that $|y| \leq 1$.

Fix $|y| \leq 1$. The roots of (1) are then $1 - \sqrt{1 - y^2}$ and $1 + \sqrt{1 - y^2}$, which are real and nonnegative. As $a_1(-x, y) = a_1(x, y)$, the interval of convergence is symmetric about $x = 0$. We shall assume then that $x \geq 0$; thus, $a_n(x, y) \geq 0$, for all n .

If $r_0 = 1 \pm \sqrt{1 - y^2}$, then $a_{n+1}(x, y)$ is less than, equal to, or greater than r_0 according to whether $a_n(x, y)$ is less than, equal to, or greater than $r_0 (= (r_0^2 + y^2)/2)$.

If $a_n(x, y)$ lies in the closed interval $[1 - \sqrt{1 - y^2}, 1 + \sqrt{1 - y^2}]$, that is, between the roots of (1), then

$$a_n(x, y)^2 - 2a_n(x, y) + y^2 \leq 0,$$

so that

$$1 - \sqrt{1 - y^2} \leq a_{n+1}(x, y) \leq a_n(x, y).$$

It follows that $(a_n(x, y))_{n \geq 0}$ converges if x is in the closed interval $[1 - \sqrt{1 - y^2}, 1 + \sqrt{1 - y^2}]$.

If $a_n(x, y)$ does not lie in the interval $[1 - \sqrt{1 - y^2}, 1 + \sqrt{1 - y^2}]$, then

$$a_n(x, y)^2 - 2a_n(x, y) + y^2 > 0,$$

so that

$$a_{n+1}(x, y) > a_n(x, y).$$

Thus, if x , and therefore all $a_n(x, y)$, are greater than $1 + \sqrt{1 - y^2}$, then the sequence diverges. On the other hand, if x , and therefore all $a_n(x, y)$, lie between 0 and $1 - \sqrt{1 - y^2}$, the sequence converges monotonically to $1 - \sqrt{1 - y^2}$.

To summarize, $(a_n(x, y))_{n \geq 0}$ converges if and only if

$$-1 \leq y \leq 1$$

and

$$-(1 + \sqrt{1 - y^2}) \leq x \leq 1 + \sqrt{1 - y^2}.$$

B-4 (35, 11, 13, 0, 0, 0, 0, 12, 5, 48, 79)

Solution. The smallest possible degree of $f(x)$ is 3984.

By the Division Algorithm, we can write $p(x) = (x^3 - x)q(x) + r(x)$, where $q(x)$ and $r(x)$ are polynomials, the degree of $r(x)$ is less than 3, and the degree of $q(x)$ is less than 1989. Then

$$\frac{d^{1992}}{dx^{1992}} \left(\frac{p(x)}{x^3 - x} \right) = \frac{d^{1992}}{dx^{1992}} \left(\frac{r(x)}{x^3 - x} \right).$$

Now, write $r(x)/(x^3 - x)$ in the form

$$\frac{A}{x - 1} + \frac{B}{x} + \frac{C}{x + 1}.$$

Because $p(x)$ and $x^3 - x$ have no nonconstant common factor, neither do $r(x)$ and $x^3 - x$, and therefore, $ABC \neq 0$. Thus,

$$\begin{aligned} & \frac{d^{1992}}{dx^{1992}} \left(\frac{r(x)}{x^3 - x} \right) \\ &= 1992! \left(\frac{A}{(x - 1)^{1993}} + \frac{B}{x^{1993}} + \frac{C}{(x + 1)^{1993}} \right) \\ &= 1992! \left(\frac{Ax^{1993}(x + 1)^{1993} + B(x - 1)^{1993}(x + 1)^{1993} + C(x - 1)^{1993}x^{1993}}{(x^3 - x)^{1993}} \right). \end{aligned}$$

Since $ABC \neq 0$, it is clear that the numerator and denominator have no common factor. Expanding the numerator yields an expression of the form

$$(A + B + C)x^{3986} + 1993(A - C)x^{3985} + 1993(996A - B + 996C)x^{3984} + \cdots.$$

From $A = C = 1$, $B = -2$, we see the degree can be as low as 3984. A lower degree would imply $A + B + C = 0$, $A - C = 0$, $996A - B + 996C = 0$, implying that $A = B = C = 0$, a contradiction.

B-5 (62, 4, 4, 0, 0, 0, 0, 3, 6, 2, 49, 73)

Solution 1. The set $\{D_n/n!\}_{n \geq 2}$ forms a sequence which strictly increases to infinity; it is therefore unbounded.

Observing that $D_2 = 3$ and $D_3 = 11$, we obtain a recursion for D_{n+1} . Subtracting the next-to-last column from the last column and then the next-to-last row from the last row, one finds

$$D_{n+1} = \det \begin{pmatrix} 3 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 4 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 5 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & n+1 & -n \\ 0 & 0 & 0 & \cdots & 0 & -n & 2n+1 \end{pmatrix}.$$

Expanding the determinant in its last row, one obtains

$$D_{n+1} = (2n+1)D_n - n^2D_{n-1}.$$

Letting $r_n = (D_n/n!)$, the recursion may be written as

$$r_{n+1} = \frac{2n+1}{n+1}r_n - \frac{n}{n+1}r_{n-1},$$

or

$$(r_{n+1} - r_n) = \frac{n}{n+1}(r_n - r_{n-1}).$$

We conclude that

$$r_{n+1} - r_n = \frac{3}{n+1}(r_3 - r_2) = \frac{1}{n+1}.$$

Therefore,

$$\begin{aligned} r_{n+1} &= r_2 + (r_3 - r_2) + (r_4 - r_3) + \cdots + (r_{n+1} - r_n) \\ &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n+1}, \end{aligned}$$

so the sequence (r_n) diverges to infinity.

Solution 2. The problem is the case $a_i = i + 1$ of

$$D_{n+1}(a_1, \dots, a_n) = \det \begin{pmatrix} 1+a_1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1+a_3 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1+a_4 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & \cdots & 1+a_n \end{pmatrix}$$

$$= \prod_{i=1}^n a_i + \sum_{i=1}^n \prod_{j=1, j \neq i}^n a_j.$$

This formula follows immediately from the recurrence

$$D_{n+1}(a_1, \dots, a_n) = a_n D_n(a_1, \dots, a_{n-1}) + a_{n-1} D_n(a_1, \dots, a_{n-2}, 0).$$

To prove this recurrence, subtract the $(n-1)$ st column from the n th column, and then expand along the n th column.

If none of the a_i 's equal 0, we can write the polynomial $D_n(a_1, \dots, a_{n-1})$ in the form

$$D_n(a_1, \dots, a_{n-1}) = a_1 a_2 \cdots a_{n-1} \left(1 + \frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_{n-1}} \right).$$

It follows that

$$\frac{D_n}{n!} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n},$$

so the sequence $(D_n/n!)$ is unbounded.

B-6 (0, 0, 0, 0, 0, 0, 0, 5, 4, 39, 155)

Solution 1. We prove the result more generally for complex matrices (because it is convenient to use $i = \sqrt{-1}$ in the proof).

The proof is by induction on n .

If $n = 1$ then the elements of \mathcal{M} commute so that (iv) cannot be satisfied unless $\mathcal{M} = \{I\}$. Suppose that $n > 1$ and that the result holds for sets of complex matrices of smaller dimension.

We may assume $|\mathcal{M}| > 1$, so by (iv), there exist $C, D \in \mathcal{M}$ with $CD = -DC$. Fix such C, D . As in the first solution, $C^2 = \pm I$. Hence the eigenvalues of C are $\pm \lambda$ where $\lambda = 1$ or i . Furthermore, $C^n = V_\lambda \oplus V_{-\lambda}$, where $V_\lambda, V_{-\lambda}$ are the nullspaces of $(C - \lambda I), (C + \lambda I)$ respectively. We observe that if $X \in \mathcal{M}$ then

$$CX = XC \Rightarrow (C \pm \lambda I)X = X(C \pm \lambda I) \Rightarrow V_{\pm \lambda} X = V_{\pm \lambda};$$

$$CX = -XC \Rightarrow (C \pm \lambda I)X = (-1)X(C \mp \lambda I) \Rightarrow V_{\pm \lambda} X = V_{\mp \lambda}.$$

In particular, since $V_\lambda D = V_{-\lambda}$, $\dim(V_\lambda) = \dim(V_{-\lambda}) = n/2$.

Let $\mathcal{N} = \{X \in \mathcal{M} \mid CX = XC, DX = XD\}$. If $Y \in \mathcal{M}$ then exactly one of Y, YC, YD, YCD is in \mathcal{N} . It follows that $|\mathcal{N}| = |\mathcal{M}|/4$.

For $X \in \mathcal{N}$, let $\phi(X)$ be the $n/2 \times n/2$ matrix representing, with respect to a fixed basis of V_λ , the linear transformation given by $v \rightarrow vX$ for $v \in V_\lambda$. Then ϕ is injective. To see this: assume $\phi(X) = \phi(Y)$ so that $vX = vY$ for $v \in V_\lambda$; but if $v \in V_{-\lambda}$ then $vD \in V_\lambda$, so that $vXD = vDX = vDY = vYD$, which again implies

$vX = vY$; since X, Y induce the same transformations of both V_λ and $V_{-\lambda}$, it follows that $X = Y$.

It suffices finally to show that $\phi(\mathcal{N})$, a set of $n/2 \times n/2$ complex matrices, satisfies (i), (ii), (iii), (iv), for then, by induction, $|\phi(\mathcal{N})| \leq (n/2)^2$, whence $|\mathcal{M}| = 4|\mathcal{N}| = 4|\phi(\mathcal{N})| \leq n^2$.

Conditions (i), (ii), (iii) for $\phi(\mathcal{N})$ are clearly inherited from those of \mathcal{M} . To show (iv), let $\phi(A) \in \phi(\mathcal{N})$, with $\phi(A)$ not the $n/2 \times n/2$ identity matrix. Then $A \neq I$ (as ϕ is injective) and $AB = -BA$ for some $B \in \mathcal{M}$. Let B' be the element of $\{B, BC, BD, BCD\}$ belonging to \mathcal{N} . Since $AB' = -B'A$, $\phi(A)\phi(B') = -\phi(B')\phi(A)$.

Solution 2. Let G be the group $\{\pm A \mid A \in \mathcal{M}\}$. We must show that $|G| \leq 2n^2$.

The center of G , $Z(G)$, consists of $\pm I$, and if $X \in G \setminus Z(G)$, then X has precisely two conjugates, namely itself and $-X$. Thus G has $1 + |G|/2$ conjugacy classes, and therefore, G has $1 + |G|/2$ inequivalent irreducible representations over \mathbb{C} .

The number of inequivalent representations of dimension 1 is $|G/G'|$, where G' is the commutator subgroup. Since $G' = \{\pm I\} = Z(G)$, this number is $|G|/2$.

The remaining irreducible representation then has dimension $\sqrt{|G|/2}$ (since the sum of the squares of the dimensions of the irreducible representations is $|G|$). This representation must be contained in the given representation of G in $n \times n$ matrices, for in all the 1-dimensional representations, $Z(G)$ is in the kernel. Hence $n \geq \sqrt{|G|/2}$, or $2n^2 \geq |G|$.

Klosinski:

*Department of Mathematics
Santa Clara University
Santa Clara, CA 95053*

Alexanderson:

*Department of Mathematics
Santa Clara University
Santa Clara, CA 95053*

Larson:

*Department of Mathematics
St. Olaf College
Northfield, MN 55057*

Professor H. B. FINE, of Princeton University, was fatally injured by an automobile on the evening of Friday, December 21 and died about one A.M. on December 22, 1928. He was seventy years of age.

36(1929), 118

A Visual Explanation of Jensen's Inequality

Tristan Needham

“This theorem is so fundamental that we propose to give a number of proofs, of varying degrees of simplicity and generality.” So say Hardy, Littlewood, and Pólya ([1], p. 17) of the theorem of the arithmetic and geometric means. True to their word, they proceed to give eleven (!) different proofs of the fact that for non-negative x_i ,

$$\sqrt[n]{x_1 \cdot x_2 \cdots x_n} \leq \left(\frac{x_1 + x_2 + \cdots + x_n}{n} \right), \quad (1)$$

with equality *iff* $x_1 = x_2 = \cdots = x_n$. For elegant applications (suitable for the classroom) of this result to elementary geometry, see [2].

One of the simplest proofs of (1) consists in recognizing it to be merely a special case of Jensen's inequality [3]. This widely used result (e.g., probability theory [4]) states that if the graph of a real continuous function $f(x)$ is *concave down* then

$$\frac{\sum f(x_i)}{n} \leq f\left(\frac{\sum x_i}{n}\right), \quad (2)$$

with equality *iff* the x 's are all equal. If the graph is concave up, the inequality is reversed. To obtain (1) we need only put $f(x) = \ln x$ and note that its graph is concave down. Very neat, but where did (2) come from? This note describes a particularly simple way of *seeing* its truth, which we hope may be of value in the classroom. Indeed, we believe it could even be used successfully in high schools.

We have given no formal definition of a graph being “concave down,” and when presenting the following argument to young students we shall suppose that none *will* be given; what matters is that they know what one looks like. With more mature students we may define the graph of f to be “concave down” if the region $\{(x, y): y \leq f(x)\}$ below the graph is convex. This is not one of the standard definitions, but it is a visually compelling inference from any other reasonable definition.

Consider a set of n point particles in the plane, of equal mass and with position vectors \mathbf{r}_i . The center of mass therefore has position vector

$$\mathbf{c} = \frac{1}{n} \sum \mathbf{r}_i,$$

from which it follows easily that

$$\sum (\mathbf{r}_i - \mathbf{c}) = \mathbf{0}.$$

In other words (see FIGURE 1), *the vectors from c to the particles cancel.*

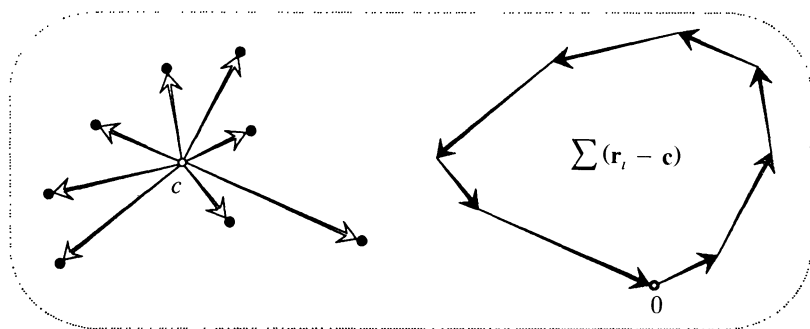


Figure 1

Imagining pegs sticking out of the plane at the locations of the particles, stretch a rubber band so as to enclose all the pegs. When released, the rubber band will contract into the dashed polygon H of FIGURE 2. This is the “convex hull” of the set of particles. The key point is this: c must lie in the shaded interior of H . For if p is outside this set, we see that the vectors from p to the particles cannot possibly cancel, as they must do for c . More formally, we take it as visually evident that through any exterior point p we may draw a line L such that H and its shaded interior lie entirely on one side of L . [Alternatively, this property may be taken as a (non-standard) *definition* of a convexity for a closed planar set.] The impossibility of the vectors cancelling now follows from their lying entirely on this side of L , for they all must have positive components in the direction of the normal vector \mathbf{n} . Except when the particles are collinear (in which case H collapses to a line-segment), the same reasoning forbids c from lying on H .

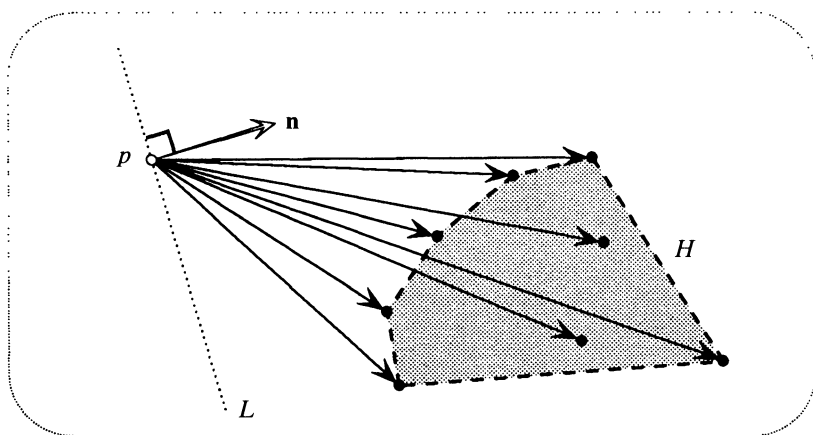


Figure 2

Next, suppose that the particles are distributed along a convex curve K . See FIGURE 3. The shaded interior of H now lies entirely on the concave side of K , and consequently so too must c . Furthermore, we see that c can only lie on K in the degenerate case that all the particles coalesce. Finally, take K to be the graph of a function $f(x)$. If this graph is *concave down* [*up*], then c lies *below* [*above*] K . Thus, with the particles located at $(x_i, f[x_i])$, we conclude that if the graph is

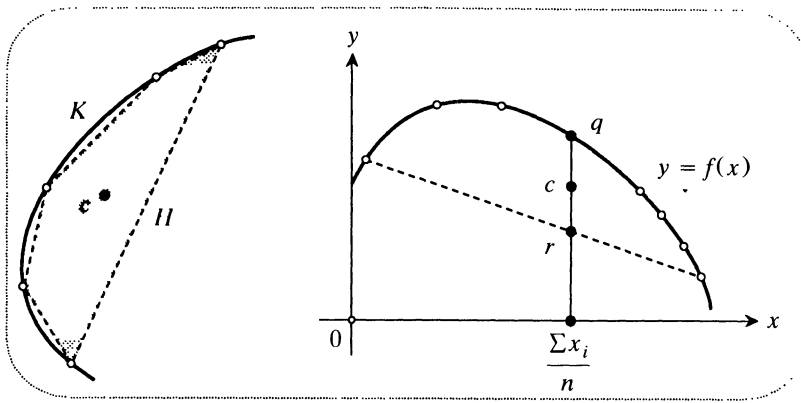


Figure 3

concave down,

$$\frac{\sum f(x_i)}{n} = \text{height of } c \leq \text{height of } q = f\left(\frac{\sum x_i}{n}\right),$$

with equality *iff* the x 's are all equal. If the graph is concave up, the inequality is simply reversed.

As a bonus, observe that c must also lie on or above the dashed chord connecting the two end particles. Thus if $y = g(x)$ is the equation of this chord, we obtain

$$g\left(\frac{\sum x_i}{n}\right) = \text{height of } r \leq \text{height of } c = \frac{\sum f(x_i)}{n}.$$

I do not know if this result has a name.

We note that the above ideas can be generalized in at least two directions:

(1) The positive masses m_i of the particles need not be equal for the argument to work. Thus, once again taking the graph to be concave down,

$$\frac{\sum m_i f(x_i)}{M} \leq f\left(\frac{\sum m_i x_i}{M}\right),$$

where M denotes the total mass. This is essentially the form that is used in probability theory, for we are free to interpret (m_i/M) as a probability distribution for x_i , yielding

$$\mathcal{E}[f(x)] \leq f(\mathcal{E}[x]),$$

where \mathcal{E} stands for the expected value. Also, by allowing the number of particles to increase without limit, we may pass from a discrete probability distribution to a continuous one.

(2) The argument is equally applicable to a set of particles in three-dimensional space. Thus, taking these particles (of equal mass, say) to be distributed over a surface $z = f(x, y)$ that is concave down, we deduce that

$$\frac{\sum f(x_i, y_i)}{n} \leq f\left(\frac{\sum x_i}{n}, \frac{\sum y_i}{n}\right).$$

Of course this too may be generalized to unequal masses and be given a probabilistic interpretation.

I do not wish to claim that the above is more original than it really is. In particular, the argument associated with FIGURE 2 is very old; I merely rediscovered it. The first important application of this idea that I know of occurred in 1874 when F. Lucas used it (see [5]) to demonstrate a complex analogue of Rolle's theorem: *the critical points of a polynomial in the complex plane must all lie within the convex hull of its zeros*. This follows from FIGURE 2 by observing [Gauss, 1816] that if $P(z)$ is the factorized polynomial, the conjugate of the logarithmic derivative $[P'(z)/P(z)]$ is a weighted sum of vectors from z to the zeros.

Also, consideration of centers of mass is certainly not new in the context of Jensen's inequality, and thus it is hard to believe that so simple a line of thought can have escaped notice. Nevertheless, it would appear that in the literature (e.g., [1], p. 71) the location of the center of mass is merely used as an *interpretation* of (2), rather than as the source of an explanation.

ACKNOWLEDGMENTS. The idea of this note arose from a conversation (several years ago) with my friend Dr. George Burnett-Stuart. I also thank Dr. John Kao for explaining to me the significance of Jensen's inequality in probability theory. Finally, I am grateful to the referee for helpful comments on the first draft.

REFERENCES

1. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 1952.
2. G. Pólya, *Mathematics and Plausible Reasoning*, vol. 1, Princeton University Press, 1954, pp. 137–141.
3. D. S. Mitrinović, *Elementary Inequalities*, P. Noordhoff Ltd., Netherlands, 1964, pp. 23–30.
4. W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd edition, Vol. 2, Wiley, 1967.
5. J. L. Walsh, *The Location of Critical Points*, Amer. Math. Soc. Colloquium Publications, Vol. XXXIV, 1950, pp. 5–7.

Mathematics Department
University of San Francisco
2130 Fulton Street
San Francisco, CA 94117-1080

The Fine Memorial Mathematics Hall, which will be erected at Princeton University at a cost of \$400,000 in memory of the late Henry B. Fine, for many years a professor of mathematics and dean of science, will be started in the near future.

36(1929), 453

The Index of a Constrained Critical Point

Catherine Hassell and Elmer Rees

1. INTRODUCTION. This note deals with the problem of determining the type of a critical point arising in the method of Lagrange multipliers. This method is the usual one used to solve the following problem:

To find the critical points of a smooth function f defined on $M^n \subset \mathbb{R}^{n+m}$, a smooth submanifold given as the common zero-set of m smooth functions $g_i: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$.

The method consists of introducing a vector $\underline{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)$ of 'undetermined multipliers', defining L to be $f + \underline{\lambda} \cdot \underline{g} = f + \sum_{i=1}^m \lambda_i g_i$ and finding its critical points. The question of deciding the non-degeneracy and type of a critical point is usually disregarded in the text books or else dismissed as being too complicated. Our purpose is to show, on the contrary, that criteria can be stated and derived in a straightforward manner.

We compare the Hessian of f restricted to M with the bordered Hessian, that is, the Hessian of L regarded as a function of $n + 2m$ variables (including $\underline{\lambda}$). The two Hessians have the same nullity at corresponding critical points and when they are non-degenerate, they have the same signature.

2. LAGRANGE MULTIPLIERS AND THE BORDERED HESSIAN. Let $U \subset \mathbb{R}^{n+m}$ be an open subset and $g: U \rightarrow \mathbb{R}^m$ be a C^1 -function such that $Dg(\underline{a}): \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ has rank m for every $\underline{a} \in M = \{\underline{x} \in U | g(\underline{x}) = \underline{c}\}$. Hence, by the implicit function theorem [F, p. 117], M is a smooth n -dimensional manifold. We wish to determine the critical points of the function $f_1: M \rightarrow \mathbb{R}$ which is the restriction of a C^2 -function $f: U \rightarrow \mathbb{R}$.

For $\underline{\lambda} \in \mathbb{R}^m$, we consider the Lagrangian

$$L = f + \underline{\lambda} \cdot (\underline{g} - \underline{c})$$

either as a function of $\underline{x} \in U$ or as a function of $(\underline{x}, \underline{\lambda}) \in U \times \mathbb{R}^m$. The critical points are obtained by solving the equations

$$\nabla L = 0 \quad \text{and} \quad \underline{g} = \underline{c}$$

or, equivalently

$$\nabla L = 0$$

regarding L as a function of $(\underline{x}, \underline{\lambda})$.

To determine the nature of a critical point \underline{a} of f_1 one could study the Taylor series of f_1 at \underline{a} in terms of local coordinates on M . Let $H_M f(\underline{a})$ be the Hessian form of f_1 ; it is the symmetric bilinear form on the tangent space $T_{\underline{a}}(M)$ which represents the quadratic terms in the Taylor expansion of f_1 . If x_1, \dots, x_n are local coordinates on M near \underline{a} , the entries of the matrix of $H_M f(\underline{a})$ are $(\partial^2 f / \partial x_i \partial x_j)$

evaluated at \underline{a} . If this matrix is non-singular, the form $H_M f(\underline{a})$ is called non-degenerate and f_1 is called a Morse function at \underline{a} . In this case the nature of the critical point is determined by the algebraic properties of $H_M f(\underline{a})$. The index of the critical point, that is the number of independent directions in which f_1 decreases, is determined by the signature of the form $H_M f(\underline{a})$. We will give practical methods for determining when $H_M f(\underline{a})$ is non-degenerate and for calculating its signature.

Let g be a C^2 -function and let $HL(\underline{a}, \underline{\lambda})$ be the bordered Hessian of L at the critical point $(\underline{a}, \underline{\lambda})$ of L ; that is, the Hessian of L regarded as a bilinear form on $T_{(\underline{a}, \underline{\lambda})}(U \times \mathbb{R}^m) \cong \mathbb{R}^{n+2m}$. If

$$\underline{g}^T = (g_1, g_2, \dots, g_m)$$

let $Dg^T = (\nabla g_1, \nabla g_2, \dots, \nabla g_m)$ denote the (transposed) Jacobian matrix of g at \underline{a} . Then the matrix of the bordered Hessian $HL(\underline{a}, \underline{\lambda})$ is the $(n + 2m)$ by $(n + 2m)$ symmetric matrix

$$\begin{bmatrix} Hf + \underline{\lambda} \cdot H\underline{g} & D\underline{g}^T \\ D\underline{g} & 0 \end{bmatrix}$$

where $\underline{\lambda}$ is evaluated from the equation

$$0 = \nabla L = \nabla f + \underline{\lambda} \cdot D\underline{g}$$

at \underline{a} .

3. THE MAIN RESULT. If a symmetric bilinear form on a real vector space is represented by the matrix H ; then its nullity is the dimension of the kernel of H and its signature is $p - q$, where p and q denote the number of positive and negative eigenvalues of H respectively.

Theorem 1. *The nullity of $H_M f(\underline{a})$ equals the nullity of $HL(\underline{a}, \underline{\lambda})$.*

If $HL(\underline{a}, \underline{\lambda})$ (and hence $H_M f(\underline{a})$) is non-degenerate, then the signature of $H_M f(\underline{a})$ equals that of $HL(\underline{a}, \underline{\lambda})$.

This theorem follows from the purely algebraic Theorem 2, using Taylor's formula and the implicit function theorem [F]. When it is applied to critical points as above, it yields the following result.

Corollary. *The point $\underline{a} \in M$ is a critical point of $f_1: M \rightarrow \mathbb{R}$ if and only if $(\underline{a}, \underline{\lambda})$ is a critical point of $L: U \times \mathbb{R}^m \rightarrow \mathbb{R}$. In this case, \underline{a} is non-degenerate if and only if $(\underline{a}, \underline{\lambda})$ is non-degenerate and the index $I(f_1, \underline{a})$ of f_1 at \underline{a} is related to the index $I(L, \underline{a}, \underline{\lambda})$ of L at $(\underline{a}, \underline{\lambda})$ by*

$$I(f_1, \underline{a}) + m = I(L, \underline{a}, \underline{\lambda}).$$

So, for example, \underline{a} is a local minimum of f_1 if $(\underline{a}, \underline{\lambda})$ is a non-degenerate critical point of L of index m .

Similarly, \underline{a} is a local maximum of f_1 if $(\underline{a}, \underline{\lambda})$ is a non-degenerate critical point of L of index $n + m$.

Theorem 2. *Let $C = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}$ be the symmetric real matrix consisting of the $(n + m) \times (n + m)$ symmetric matrix A , the $m \times (n + m)$ matrix B of rank m and the zero*

$m \times m$ matrix 0. The symmetric bilinear form induced on $\text{Ker } B$ by A is denoted b . Then the bilinear form on \mathbb{R}^{n+2m} defined by C is isomorphic to $b \oplus H$ where H is the $2m$ -dimensional hyperbolic form $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$.

The proof that we give for this theorem is considerably simpler than our original one and is based on a proof provided for us by Dr. A. A. Ranicki.

First, we give proofs of some facts from linear algebra that we need.

Fact 1 (Fredholm alternative) [HK, p. 103]. Let $P^T: \mathbb{R}^k \rightarrow \mathbb{R}^l$ be the transpose of the matrix $P: \mathbb{R}^l \rightarrow \mathbb{R}^k$, then

$$(\text{Ker } P)^\perp = \text{Im } P^T.$$

Proof: Let r denote rank P . Then $\dim \text{Ker } P = l - r$ hence $\dim(\text{Ker } P)^\perp = r$. Also $\dim \text{Im } P^T = r$. It is therefore enough to show that $\text{Im } P^T \subset (\text{Ker } P)^\perp$ i.e.

$$\text{Im } P^T \perp \text{Ker } P.$$

Suppose $\underline{x} \in \mathbb{R}^k$ and $\underline{z} \in \text{Ker } P$ then

$$\underline{z} \cdot P^T \underline{x} = \underline{z}^T P^T \underline{x} = 0 \quad \text{since } P \underline{z} = 0.$$

Let b be a bilinear form on the real finite dimensional space V , the annihilator of a subspace $U \subset V$ is $U^\perp = \{\underline{x} \in V \mid b(\underline{x}, \underline{u}) = 0 \ \forall \underline{u} \in U\}$.

Fact 2. Let b be a non-degenerate symmetric bilinear form on V of dimension $2m$ and let $W \subset V$ have dimension m and $W \subset W^\perp$. Then b is represented by the matrix $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$.

Proof: Choose $\underline{w} \neq 0$ in W and \underline{v} such that $b(\underline{w}, \underline{v}) = 1$. Let $\underline{u} = \underline{v} - b(\underline{v}, \underline{v})\underline{w}/2$, then $b(\underline{u}, \underline{u}) = 0$ and $\{\underline{u}, \underline{w}\}$ is the required basis in the case $m = 1$. When $m > 1$, let $U = \text{Span}\{\underline{u}, \underline{w}\}$ and consider U^\perp , this contains a subspace that is self-annihilating and of dimension $m - 1$. The result follows by induction.

We also make use of the principal axis theorem.

Fact 3 [HK, p. 266]. If b is a symmetric bilinear form on an n -dimensional real inner product space then there is an orthonormal basis $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ such that $b(\underline{e}_i, \underline{e}_j) = \delta_{ij}a_i$ for some $a_i \in \mathbb{R}$.

Proof of Theorem 2: If W denotes the m -dimensional subspace

$$\left\{ \begin{pmatrix} 0 \\ \underline{y} \end{pmatrix} : \underline{y} \in \mathbb{R}^m \right\} \subset \mathbb{R}^{n+2m},$$

then by Fact 1 applied to B , there is a canonical orthogonal decomposition

$$\mathbb{R}^{n+2m} \cong \text{Ker } B \oplus \text{Im } B^T \oplus W.$$

By Fact 3, one can choose an orthonormal basis

$$\{\underline{e}_1, \dots, \underline{e}_n\} \text{ for } \text{Ker } B$$

such that $b(\underline{e}_i, \underline{e}_i) = a_i$ and $b(\underline{e}_i, \underline{e}_j) = 0$ for $i \neq j$. Then for $1 \leq i \leq n$, \underline{e}_i is a vector whose component in W is zero. Since $\underline{e}_i \in \text{Ker } B$ one has $C\underline{e}_i \in W^\perp$ and hence $C\underline{e}_i = \underline{k}_i + B^T \underline{f}_i$ where $\underline{k}_i \in \text{Ker } B$ and $\underline{f}_i \in \mathbb{R}^m \cong W$ is unique because B^T

is one-to-one. Moreover, $\underline{k}_i = \underline{a}_i \underline{e}_i$ because $\underline{e}_j^T C \underline{e}_i = \delta_{ij} a_i$. Hence,

$$\begin{aligned} C \underline{e}_i &= a_i \underline{e}_i + B^T \underline{f}_i \\ &= a_i \underline{e}_i + C \underline{f}_i \end{aligned}$$

Define $K = \text{Span}\{\underline{e}_i - \underline{f}_i; 1 \leq i \leq n\}$.

The following steps will prove Theorem 2.

Step 1. K and $\text{Im } B^T \oplus W$ are orthogonal with respect to C .

Step 2. The form defined by C on K is isomorphic to b .

Step 3. The form defined by C on $\text{Im } B^T \oplus W$ is hyperbolic.

Proof of Step 1: Since $C(\underline{e}_i - \underline{f}_i) = a_i \underline{e}_i \in \text{Ker } B$, one has that $C(K) \subset \text{Ker } B$ and $\text{Ker } B$ is orthogonal to $\text{Im } B^T \oplus W$.

Proof of Step 2: Since $(\underline{e}_j - \underline{f}_j)^T C(\underline{e}_i - \underline{f}_i) = a_i \delta_{ij}$, one has that $C|_K$ is isomorphic to b .

Proof of Step 3: By choosing a basis for the image of B^T , one can take the matrix of C to have the following form

$$\begin{bmatrix} A_1 & A_2^T & 0 \\ A_2 & A_3 & I_m \\ 0 & I_m & 0 \end{bmatrix}.$$

Hence the matrix of $C|_{\text{Im } B^T \oplus W}$ has the form

$$\begin{bmatrix} A_3 & I_m \\ I_m & 0 \end{bmatrix}$$

and so is equivalent to a hyperbolic form by Fact 2, since W is a self-annihilating subspace of dimension m .

4. COMPARISON WITH CLASSICAL CRITERIA. In the literature there are criteria for deciding when a critical point is a local maximum or minimum, for example [H] or [G]. Here we show how these criteria are related to our result.

Criterion 1. Let $C = \begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}$ be as in Theorem 2 and assume that the last $m \times m$ submatrix of B is non-singular, then the form induced by A on $\text{Ker } B$ is positive definite if the determinants Δ_i for $0 \leq i \leq n$ have sign $(-1)^m$ where $\Delta_i = \det C_i$ and C_i is obtained from C by deleting its first i rows and columns.

Proof: Write $C_n = \begin{bmatrix} A_n & B_n^T \\ B_n & 0 \end{bmatrix}$. Then $\Delta_n = \det C_n = (-1)^m (\det B_n)^2$, so $\text{sign } \Delta_n = (-1)^m$ since B_n is non-singular. By Fact 2, C_n is hyperbolic and so has index m .

The proof is completed by using induction based on the following:

Lemma. Let H be a non-singular symmetric real matrix and H_1 be obtained from H by deleting one row and the corresponding column. If H_1 is also non-singular and

index H is the number of negative eigenvalues of H then

$$\text{index } H = \begin{cases} \text{index } H_1 \\ \text{index } H_1 + 1 \end{cases}$$

depending on whether $\det H$ and $\det H_1$ have the same or the opposite sign.

Proof: Let M and M_1 be maximal negative definite subspaces for H and H_1 respectively.

Recall that the dimension of a maximal negative definite subspace is unique.

Clearly, $\dim M_1 \leq \dim M \leq \dim M_1 + 1$.

Also $\text{sign } \det H = (-1)^{\dim M}$ and $\text{sign } \det H_1 = (-1)^{\dim M_1}$.

Hence $\det H$ and $\det H_1$ have the same sign $\Leftrightarrow \dim M = \dim M_1$ as required.

Another criterion discovered in the 19th century is the following (see [H] for the historical references).

Criterion 2. Let $\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix}$ be as in Theorem 2. Then the form induced by A on $\text{Ker } B$ is positive definite if and only if the roots of

$$\det \begin{bmatrix} A - tI & B^T \\ B & 0 \end{bmatrix} = 0$$

are all positive.

Note that the above equation is of degree n .

The stronger result that the roots of the above equation are the eigenvalues of the form A restricted to $\text{Ker } B$ with the same multiplicities is an immediate consequence of Theorem 2 applied to the matrix

$$\begin{bmatrix} A - tI & B^T \\ B & 0 \end{bmatrix}$$

when t is a root of the above equation.

5. EXAMPLES

1. To find the critical points of $f(x, y, z) = x^3 + y^3 + z^3$ on the surface $x^{-1} + y^{-1} + z^{-1} = 1$. (This example is taken from [G, p. 94]).

Let $L = x^3 + y^3 + z^3 + \lambda(x^{-1} + y^{-1} + z^{-1} - 1)$ then $(\partial L / \partial x) = 3x^2 - \lambda x^{-2}$ etc. and the bordered Hessian is

$$\begin{bmatrix} 6x + 2\lambda x^{-3} & 0 & 0 & -x^{-2} \\ 0 & 6y + 2\lambda y^{-3} & 0 & -y^{-2} \\ 0 & 0 & 6z + 2\lambda z^{-3} & -z^{-2} \\ -x^{-2} & -y^{-2} & -z^{-2} & 0 \end{bmatrix}.$$

The critical points are given by

$$x^4 = y^4 = z^4 = \lambda/3 \text{ and } x^{-1} + y^{-1} + z^{-1} = 1.$$

These are $x = y = z = 3$, $\lambda = 243$; $x = y = 1$, $z = -1$, $\lambda = 3$ and two other solutions symmetrical with the latter.

In the first case the Hessian is

$$\begin{pmatrix} 36 & 0 & 0 & -9^{-1} \\ 0 & 36 & 0 & -9^{-1} \\ 0 & 0 & 36 & -9^{-1} \\ -9^{-1} & -9^{-1} & -9^{-1} & 0 \end{pmatrix}$$

which is non-degenerate and has signature 2, so the critical point has index 0; that is, it is a non-degenerate minimum.

In the second case the Hessian is

$$\begin{pmatrix} 12 & 0 & 0 & -1 \\ 0 & 12 & 0 & -1 \\ 0 & 0 & -12 & -1 \\ -1 & -1 & -1 & 0 \end{pmatrix}$$

which is non-degenerate and has signature 0. So this critical point (and the other two symmetrical with it) has index 1; that is, it is a saddle point.

2. Consider the quadratic form $\underline{x}^T A \underline{x}$ on the sphere $\underline{x}^T \underline{x} = 1$ in \mathbb{R}^n . Critical points \underline{x} are given by

$$A \underline{x} - \lambda \underline{x} = 0$$

i.e. by an eigenvector \underline{x} with eigenvalue λ . The critical point is non-degenerate if $\begin{bmatrix} A - \lambda I & \underline{x} \\ \underline{x}^T & 0 \end{bmatrix}$ is non-singular and this is true if the eigenvalue has multiplicity 1. If the eigenvalue has multiplicity $r > 1$, let $\underline{x}_1, \dots, \underline{x}_r$ be a basis for the eigenspace, then

$$\begin{bmatrix} A - \lambda I & \underline{x}_1 & \cdots & \underline{x}_r \\ \underline{x}_1^T & & & \\ \vdots & & \mathbf{0} & \\ \underline{x}_r^T & & & \end{bmatrix}$$

is non-singular. The corresponding critical submanifold is a great sphere of dimension $r - 1$ and is non-degenerate in Bott's sense [B]. We recall this concept briefly. Let S be a connected critical submanifold for a function f ; it is called non-degenerate if for every $\underline{x} \in S$, the Hessian is non-degenerate normal to S , i.e. the Hessian is zero on $T_{\underline{x}} S$ and the induced form on $T_{\underline{x}} M / T_{\underline{x}} S$ is non-degenerate. If $M^n \subset \mathbb{R}^{n+m}$ is defined by $\underline{g} = \underline{c}$, then the bordered Hessian is

$$HL(\underline{a}, \underline{\lambda}) = \begin{bmatrix} Hf + \underline{\lambda} \cdot H\underline{g} & D\underline{g}^T \\ D\underline{g} & 0 \end{bmatrix}.$$

If $\{\underline{e}_1, \dots, \underline{e}_l\}$ is a basis for $T_{\underline{a}} S$, $\underline{e}_i \in \mathbb{R}^{n+m}$, then S is non-degenerate at \underline{a} if

$$\begin{bmatrix} Hf + \underline{\lambda} \cdot H\underline{g} & D\underline{g}^T & E^T \\ D\underline{g} & 0 & 0 \\ E & 0 & 0 \end{bmatrix}$$

is non-singular, where E^T is the matrix $(\underline{e}_1, \underline{e}_2, \dots, \underline{e}_l)$. The signature of this enlarged bordered Hessian determines the index of the critical submanifold in the same way as our main result deals with non-degenerate critical points.

Early references that discuss the general problem are listed in [H, chapter VI]. Some of the more recent references for the problem are [M2], [D], [G], [S2], [M1]. The paper [S1] considers the case of a single constraint equation and includes the results of this note in that case.

- [B] R. Bott, Non-degenerate critical manifolds, *Ann. of Math.* (2) 60 (1954) 248–261.
- [D] G. Debreu, Definite and semidefinite quadratic form, *Econometrica* 20 (1952) 295–300.
- [F] W. H. Fleming, *Functions of Several Variables*, Addison Wesley, 1965.
- [G] R. P. Gillespie, *Partial Differentiation*, Oliver & Boyd, 1954.
- [H] H. Hancock, *Theorem of Maxima and Minima*, Ginn & Co., 1917, Dover, 1960.
- [HK] K. Hoffman and R. Kunze, *Linear Algebra*, Prentice Hall, 1961.
- [M1] J. H. Maddocks, Restricted quadratic forms, inertia theorems and the Schur complement, *Linear Alg. Appl.* 108 (1988) 1–36.
- [M2] H. B. Mann, Quadratic forms with linear constraints, *Amer. Math. Monthly* 50 (1943) 430–433.
- [S1] T. Sakalis, The computation of the index of a Morse function at a critical point, *Internat. J. Math. & Math. Sci.* 11 (1988) 721–726.
- [S2] D. Spring, On the second derivative test for constrained local extrema, *Amer. Math. Monthly* 92 (1985) 631–643.

Department of Mathematics and Statistics
The University of Edinburgh
Edinburgh EH9 3JZ, Scotland

A Characterization of Euclidean Spaces

In connection with the article “A Characterization of Inner Product Spaces” by Neil Falkner (this *Monthly* 100(1993), 246–249) it might be worth noting that inner product spaces over the reals are characterized by the validity of the Converse Theorem of Pythagoras. The latter, namely that the smaller sides of a triangle which fulfills the famous Pythagorean relation $a^2 + b^2 = c^2$ are orthogonal, is often assumed without proof as for instance in the argument about the legendary rope-stretchers of Ancient Egypt, who are said to have used a triangle with sides 3, 4, 5 to construct a right angle.

In the notions of an inner product space we have $\|x + y\|^2 = \|x\|^2 + 2\operatorname{re}(x, y) + \|y\|^2$. So the Theorem of Pythagoras and its converse are obvious in the case of a real inner product space. However, in any complex inner product space (with the exception of the trivial space $\{0\}$, which is no real complex space anyway) we may take $x \neq 0$ and $y := i x$ such that $(x, y) = i\|x\|^2 \neq 0$, but still $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ holds true.

The fact that ‘in Euclid’s *Elements* the Theorem of Pythagoras (I.47) is followed by the Converse Theorem of Pythagoras (I.48) and its proof is another justification for calling inner product spaces over the reals Euclidean spaces.

Andreas M. Hinz
 Mathematisches Institut
 Universität München
 Theresienstraße 39
 D-80333 München

andreas.hinz@mathematik.uni-muenchen.dbp.de

On Some Irrational Decimal Fractions

Norbert Hegyvári

It is known that the decimal fraction

$$\alpha = 0.235711131719 \dots$$

is irrational, where the sequence of digits is formed by the primes in ascending order. In [1, Th. 138] there are two different proofs for this statement. The first uses a special case of the Dirichlet's theorem, namely: any arithmetical progression of the form $10^{s+1}k + 1$ ($k = 1, 2, \dots$) contains primes. In the second proof it is assumed that there is a prime between N and $10N$ for every $N > 0$, which is the special case of the Bertrand's Postulate. Similar proofs are found in [2].

In this article we will give a direct proof for this statement. We prove even more.

Theorem. *Let $1 \leq a_1 < a_2 < \dots$ be a sequence of integers for which $\sum_{i=1}^{\infty} 1/a_i = \infty$. Then the decimal fraction $\alpha = 0 \cdot (a_1)(a_2) \dots (a_n) \dots$ is irrational.*

Since $\sum_{i=1}^{\infty} 1/p_i = \infty$, where $p_1 < p_2 < \dots$ is the sequence of primes, we immediately get the original version of the statement.

Definition. Let B be a block of digits $b_1 b_2 \dots b_s$ with $s \geq 1$ and $0 \leq b_i \leq 9$ for $i = 1, 2, \dots, s$. Let n be a positive integer $\sum_{i=0}^k c_i 10^{k-i}$ with $c_0 \neq 0$. The integer n is said to contain the block of digits B if for some $j \geq 0$ we have $c_{i+j} = b_i$ for every $i = 1, 2, \dots, s$. For example, the integer 1402857 contains the blocks 14 and 0285 (among others), but not the blocks 014 or 582.

Lemma. *If $X = X(b_1, b_2, \dots, b_s)$ denotes the sequence of positive integers not containing the block of digits $b_1 b_2 \dots b_s$, then $\sum_{n \in X} 1/n$ is convergent.*

We mention that the Lemma is a generalization of a well-known exercise (see [1, Th 144]).

Proof of the Lemma: Let $s_n = 1/x_1 + 1/x_2 + \dots + 1/x_n$ and let t be an integer for which $x_{t-1} < 10^s \leq x_t$. Then we have

$$s_n < 1/x_1 + 1/x_2 + \dots + 1/x_t + 10^{-s} (1/\lfloor x_{t+1}/10^s \rfloor + \dots + 1/\lfloor x_n/10^s \rfloor).$$

We note that if $t < i \leq n$, then $\lfloor x_i/10^s \rfloor$ is a member of X , say x_j . Also, since the block $b_1 b_2 \dots b_s$ appears in at least one of 10^s consecutive integers, it follows that for any fixed x_j there are at most $10^s - 1$ values of x_j such that $\lfloor x_i/10^s \rfloor = x_j$, and we have

$$s_n < \sum_{i=1}^t 1/x_i + (10^s - 1)10^{-s}s_n \quad \text{or} \quad s_n < 10^s \cdot \sum_{i=1}^t x_i,$$

which proves the lemma.

Proof of the Theorem: Assume that α is a rational number. Thus α is a periodic decimal, with a block of digits, say $b_1b_2\ldots b_s$, repeating endlessly perhaps after an initial first block. If B is a block of 1's, define $c_1c_2\ldots c_{2s}$ to be a block of 2's of length $2s$; otherwise define $c_1c_2\ldots c_{2s}$ to be a block of 1's of length $2s$. Now define $Y = Y(c_1, c_2, \ldots, c_{2s})$ as the sequence of natural numbers not containing the block of digits $c_1c_2\ldots c_{2s}$. If we write

$$\sum_{i=1}^{\infty} 1/a_i = \sum_{a \in Y} 1/a + \sum_{a \notin Y} 1/a,$$

then by the Lemma the first sum on the right side converges, and hence the second sum diverges. This implies that there are infinitely many a_i that contain the block of digits $c_1c_2\ldots c_{2s}$. This in turn implies that B cannot be a repeating block of digits in α . This contradiction establishes the Theorem.

ACKNOWLEDGMENT. The author would like to thank the referee for a number of suggestions and for detecting some flaws in our original version.

REFERENCES

1. Hardy–Wright, *An Introduction to the Theory of Numbers*, fifth edition, Oxford, Clarendon Press, 1979.
2. G. Pólya–G. Szegő, *Problems and Theorems in Analysis II.*, Springer-Verlag, 1976, (exercise 257.)

*Department of Math. L. Eötvös Univ. and
Math. Inst. of the Hung. Acad. of Sci.
Budapest, Pf 127, H-1364
Hungary*

Professor Florian Cajori died suddenly of pneumonia on August 14, 1930, at his home in Berkeley, California. He was a charter member of the Mathematical Association of America and was one of an original group of four (later enlarged to twelve) representatives of mid-western universities and colleges who made possible the re-establishment of the American Mathematical Monthly on a sound financial basis. A detailed account of his historical researches will be published in the *Monthly* in due course.

37(1930), 392

NOTES

Edited by: John Duncan

The Symmetry Principle for Möbius Transformations

Louis Brickman

With precise definitions to come below, the symmetry principle is the following.

Theorem. *Let E be a circle or extended line. Let T be a Möbius transformation. Let z and z^* be symmetric points with respect to E . Then $T(z)$ and $T(z^*)$ are symmetric with respect to $T(E)$ (which is also a circle or extended line).*

Can the discussion be both rigorous and intuitively satisfying? The key is that the theorem is more about “conjugate Möbius transformations” than ordinary Möbius transformations. Indeed, the theorem holds for either type of transformation, whereas the symmetry concept involves only the former. To prepare the proof we need only set down the composition relationships between the two types (Lemma 1), and then show that each circle or extended line determines a unique and very special conjugate Möbius transformation (Lemma 2).

Many of the standard proofs establish the conclusion separately for special transformations such as translations, inversions, and dilations. The results are then combined in a composition argument. Another well known approach depends upon the concept of cross ratio. The method here seems simplest.

Preliminary Definitions. The complex plane and extended complex plane are denoted by \mathbf{C} and $\hat{\mathbf{C}}$, respectively ($\hat{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$). A Möbius transformation is a map $T: \hat{\mathbf{C}} \rightarrow \hat{\mathbf{C}}$ defined by

$$T(z) = \frac{az + b}{cz + d} \quad (a, b, c, d, \in \mathbf{C}; ad - bc \neq 0).$$

The formula is extended by continuity for $z = \infty$ and, if $c \neq 0$, for $z = -d/c$. With each such T we associate the conjugate Möbius transformation $\bar{T}: \hat{\mathbf{C}} \rightarrow \hat{\mathbf{C}}$ defined by

$$\bar{T}(z) = \overline{T(\bar{z})} \quad (\bar{\infty} = \infty).$$

Finally, we let \mathcal{M} be the set (actually “group”) of all Möbius transformations and $\bar{\mathcal{M}}$ be the set of all conjugate Möbius transformations.

We remark (without proof) that our first lemma is equivalent to the statement that $\mathcal{M} \cup \bar{\mathcal{M}}$ is a group, and \mathcal{M} and $\bar{\mathcal{M}}$ are the cosets of the normal subgroup \mathcal{M} .

Lemma 1. *Let $S, T \in \mathcal{M}$. Then*

$$(1) T \circ S \in \mathcal{M}, \quad (2) T \circ \bar{S} \in \bar{\mathcal{M}}, \quad (3) \bar{T} \circ S \in \bar{\mathcal{M}}, \quad (4) \bar{T} \circ \bar{S} \in \mathcal{M}.$$

Proof: Conclusion (1) is standard. Then (3) follows immediately because $\bar{T} \circ S = \overline{T \circ S}$. Once (2) is proved, (4) follows from the fact that $\bar{T} \circ \bar{S} = \overline{T \circ S}$. Thus we need only prove (2). With T as described above,

$$(T \circ \bar{S})(z) = T(\bar{S}(z)) = \frac{a\bar{S}(z) + b}{c\bar{S}(z) + d} = \left[\frac{\bar{a}S(z) + \bar{b}}{\bar{c}S(z) + \bar{d}} \right]^{-}.$$

Conclusion (2) now follows from (1).

Lemma 2. *For each circle or extended line E , there is a unique $\bar{T} \in \mathcal{M}$ such that*

$$E = \{z \in \hat{\mathbb{C}} : \bar{T}(z) = z\}.$$

(E is exactly the set of fixed points of \bar{T} .) This \bar{T} is an involution of $\hat{\mathbb{C}}$; that is, $\bar{T} \circ \bar{T}$ is the identity.

Proof: A circle described by $|z - a| = r (r > 0)$ is equivalently described by

$$z = \bar{T}(z) = \frac{r^2}{z - a} + a.$$

A line $\{a + bt : t \in \mathbb{R}\} (b \neq 0)$ has the equation

$$\frac{z - a}{b} = \left(\frac{z - a}{b} \right)^{-}, \quad \text{or} \quad z = \bar{T}(z) = \frac{b}{\bar{b}}(\overline{z - a}) + a.$$

Since $\bar{T}(\infty) = \infty$, the extended line $\{a + bt : t \in \mathbb{R}\} \cup \{\infty\}$ is exactly the set of fixed points of \bar{T} .

For uniqueness suppose $\bar{T}_1(z) = z = \bar{T}_2(z)$ for all z on a circle or extended line. Then $T_1(z) = T_2(z)$ for more than 2 values of z . Hence $T_1 = T_2$ and $\bar{T}_1 = \bar{T}_2$. (The uniqueness of \bar{T} for an extended line may be surprising in view of the fact that a and b are not uniquely determined by the extended line.)

For the involution proof we note that $\bar{T} \circ \bar{T} \in \mathcal{M}$ (Lemma 1, part (4)) and has all the points of E as fixed points.

Definition. The transformation \bar{T} described in LEMMA 2 is called *reflection in E* , and will be denoted by ρ_E . If confusion seems unlikely, $\rho_E(z)$ is denoted simply by $z^* (z \in \hat{\mathbb{C}})$. Also, z and z^* are said to be *symmetric with respect to E* .

Now that reflection is solidly defined it is easy to prove the theorem. With obvious changes the proof applies equally well to conjugate Möbius transformations.

Proof of Symmetry Principle: In precise terms we must show that

$$\rho_{T(E)}(T(z)) = T(\rho_E(z)) \quad (z \in \hat{\mathbb{C}}),$$

or

$$\rho_{T(E)} \circ T = T \circ \rho_E.$$

But both sides of the last equation belong to \mathcal{M} (Lemma 1, parts (2) and (3)), and they agree everywhere on E . Therefore we are finished.

*Department of Mathematics and Statistics
State University of New York at Albany
Albany, New York 12222*

A Short Proof for Romberg Integration

T. von Petersdorff

The Romberg extrapolation method for numerical integration is discussed in most numerical analysis textbooks. We give a short proof for the convergence rates of the Romberg extrapolations without using the Euler-Maclaurin formula.

The Romberg method starts with the sequence of values T_N of the composite trapezoid rule with $N = 1, 2, 4, 8, \dots$ subintervals which converges to the exact integral with a rate of $O(N^{-2})$. By using linear combinations of the values T_N new sequences $T_{N,1}, T_{N,2}, \dots$ are constructed which converge to the exact integral with the rates of $O(N^{-4}), O(N^{-6}), \dots$ for $N \rightarrow \infty$. We will see that the gain of two powers of N with each extrapolation step is due to the symmetry of the trapezoid rule.

The classical proof of the Romberg method on an interval uses the Euler-Maclaurin formula to derive an asymptotic expansion of the error of the composite trapezoid rule (e.g., [2]). The convergence rate of the Romberg extrapolations then follows from this expansion and the fact that it contains only even powers of N .

The proof of the Euler-Maclaurin formula is elementary. But the proof is based on certain recursion properties of the Bernoulli polynomials and it is not intuitively obvious what it is that makes the Romberg method work.

The convergence properties of the Romberg method can be understood by using homogeneity and symmetry principles, see e.g. [1] and the references given there. Here we want to give a simple proof which only uses these two basic principles (and Taylor's theorem). We will only derive the convergence rates of the extrapolated values based on the sequence of $1, 2, 4, 8, \dots$ subintervals. We do not obtain a general asymptotic expansion or formulae for the constants in the estimates. For results of this type see [2], [1] and the references given there.

THE ROMBERG INTEGRATION METHOD. Let f be a continuous function on the interval $[a, b]$, and let $I(f) = \int_a^b f(x) dx$. The *trapezoid rule* on $[a, b]$ is defined by

$$T^{[a,b]}(f) = \frac{1}{2}(b-a)(f(a) + f(b))$$

and the *composite trapezoid rule* with N subintervals on $[a, b]$ is given by

$$T_N^{[a,b]}(f) = \sum_{k=1}^N T^{[x_{k-1}, x_k]}(f) = \frac{b-a}{2N} \left(f(a) + f(b) + 2 \sum_{k=1}^{N-1} f(x_k) \right)$$

where $x_k = a + k(b-a)/N$, $k = 0, \dots, N$. The *Romberg extrapolations* are defined recursively by

$$T_{k,0}(f) = T_k^{[a,b]}(f), \quad T_{2k,m+1}(f) = \frac{2^{2m+2}T_{2k,m}(f) - T_{k,m}(f)}{2^{2m+2} - 1}$$

for integers $k \geq 1$, $m \geq 0$. The convergence is given by the following theorem:

Theorem 1. Let f be $2m+2$ times continuously differentiable on $[a, b]$. Let $N = 2^m n$ with positive integers m, n . Then

$$|T_{N,m}(f) - I(f)| \leq C_m(b-a)^{2m+3} \max_{\xi \in [a,b]} |f^{(2m+2)}(\xi)| n^{-(2m+2)} \quad (1)$$

Here the constant C_m is independent of f, a, b, N .

THE PROOF. We consider an integration rule on $[-1, 1]$ of the form

$$A(f) = \sum_{j=1}^J f(\xi_j) w_j \quad (2)$$

with certain nodes $\xi_j \in \mathbb{R}$ and weights $w_j \in \mathbb{R}$, $j = 1, \dots, J$. The corresponding rule for the interval $[a, b]$ is given by $A^{[a,b]}(f) = \frac{1}{2}(b-a)A(\tilde{f})$ where $\tilde{f}(x) = f((a+b) + (b-a)x)/2$. We will use the following theorem which is a standard result in numerical analysis textbooks and follows from Taylor's theorem.

Theorem 2. Assume the integration rule (2) is exact for all polynomials of degree less than or equal to r , let f be $r+1$ times continuously differentiable. Then the composite rule $A_N^{[a,b]}(f) = \sum_{k=1}^N A^{[x_{k-1}, x_k]}(f)$ satisfies for all positive integers N

$$|A_N^{[a,b]}(f) - I(f)| \leq C(b-a)^{r+2} \max_{\xi \in [a,b]} |f^{(r+1)}(\xi)| N^{-(r+1)}$$

where the constant C is independent from f, a, b, N .

We now make the following assumption about $A(f)$:

Assumption 1. Let the integration rule $A(f)$ on $[-1, 1]$ be symmetric with respect to 0, i.e., $A(\tilde{f}) = A(f)$ where $\tilde{f}(x) = f(-x)$, for all continuous f . Furthermore, let the rule $A(f)$ be exact for all polynomials of degree less than or equal to q with some even number q .

Then we have

Proposition 1. The rule $A(f)$ is also exact for all polynomials of degree $q+1$.

Proof: The function x^{q+1} is odd, hence $A(x^{q+1}) = 0$ and $\int_{-1}^1 x^{q+1} dx = 0$.

Now we consider the composite rule $A_2(f(x)) = \frac{1}{2}A(f((x-1)/2)) + \frac{1}{2}A(f((x+1)/2))$ with two subintervals on $[-1, 1]$ and denote the quadrature errors by $E(f) = A(f) - I(f)$, $E_2(f) = A_2(f) - I(f)$. Then

$$E_2(x^{q+2}) = \frac{1}{2}E\left(\left(\frac{x-1}{2}\right)^{q+2}\right) + \frac{1}{2}E\left(\left(\frac{x+1}{2}\right)^{q+2}\right) = 2^{-(q+2)}E(x^{q+2}) \quad (3)$$

by expanding the powers and using Proposition 1. Therefore we can integrate x^{q+2} exactly with the rule

$$\tilde{A}(f) = \frac{2^{q+2}A_2(f) - A(f)}{2^{q+2} - 1} \quad (4)$$

Hence this construction implies:

Proposition 2. The rule $\tilde{A}(f)$ is symmetric and exact for all polynomials of degree less than or equal to $q+2$.

Now Theorem 1 follows by induction: Let $A^0(f) = T^{[-1,1]}(f)$. Obviously this rule satisfies Assumption 1 with $q = 0$. Assume that the rule $A^m(f)$ satisfies

Assumption 1 with $q = 2m$. Define the rule $A^{m+1}(f) = \widetilde{A^m}(f)$ using (4) with $q = 2m$. By Proposition 2, $A^{m+1}(f)$ satisfies Assumption 1 with $q = 2m + 2$.

By Proposition 1, the rule $A^m(f)$ is actually exact for all polynomials of degree less than or equal to $2m + 1$. Finally note that for $N = 2^m n$ we have $T_{N,m}(f) = (A^m)_n^{[a,b]}(f)$ where $(A^m)_n^{[a,b]}(f)$ denotes the composite rule on $[a, b]$ with n subintervals which is based on the rule A^m . Therefore Theorem 2 implies (1).

Note Theorem 1 remains true if we replace the trapezoid rule by any other symmetric rule.

Remark. Romberg integration on triangles can be treated in a similar way: Here the basic rule T uses the function values at the three vertices of the triangle, this rule is exact for polynomials of total degree one or less. For a rule A on a triangle we define the composite rule A_N by dividing the triangle in N^2 congruent smaller triangles and applying the basic rule A on each subtriangle. Assume that the rule A is exact for all polynomials of total degree q or less with q even. Let E and E_2 be the integration errors of A and A_2 . If f_{q+1} and f_{q+2} are monomials of total degree $q + 1$ and $q + 2$, respectively, then we obtain

$$E_2(f_{q+1}) = 2^{-(q+2)}E(f_{q+1}), \quad E_2(f_{q+2}) = 2^{-(q+2)}E(f_{q+2}). \quad (5)$$

Hence the rule \tilde{A} defined by (4) will be exact for polynomials of total degree $q + 2$ or less. To prove (5), consider the triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$. Then proceed analogously as in (3) and note that one of the four subtriangles is rotated by 180 degrees. Therefore one of the four terms arising from $E_2(f_{q+1})$ has the opposite sign. For $E_2(f_{q+2})$ expand the arising terms in monomials of degree $q + 2$, $q + 1$, and lower order terms. Then the terms of order $q + 1$ will cancel each other since the central subtriangle is rotated by 180 degrees. As the rule $A^0 = T$ is exact for polynomials of degree zero, induction shows that the rule A^m is exact for polynomials of total degree $2m$ or less. Hence the Romberg extrapolations $T_{N,m}$ converge with order $O(N^{-(2m+1)})$. This is one order lower than in the one-dimensional case, and this result cannot be improved. But no symmetry of the underlying rule T is required for this argument, so T can be *any* quadrature rule which is exact for the function 1.

REFERENCES

1. J. N. Lyness, Quadrature over a simplex: Part 2. A representation for the error functional, *SIAM J. Numer. Anal.*, 5 (1978) 870–887.
2. J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1992.

*Department of Mathematics
University of Maryland
College Park, MD 20742
tvp@math.umd.edu*

An Elementary Proof that the Borromean Rings Are Non-Splittable

Ollie Nanyes

Linström and Zetterström [1] gave a proof that the Borromean rings (figure 1) could not consist of true circles. In this note, we give an elementary proof (sans algebraic topology) that the Borromean rings are “linked” though no two components are. The tool that we use is the colorability *mod* n of a knot or link diagram. This tool has been presented in honors undergraduate seminars. I have included a discussion of colorability *mod* n though the technique is well known. For example, see Kauffman, Chapter VI [2].

1. DEFINITIONS. A *knot* will be defined as a smooth (or polyhedral) simple closed curve in 3-space R^3 . A *link* is defined as a collection of disjoint smooth (or polyhedral) simple closed curves in R^3 . Two knots or links K_1 and K_2 are said to be *equivalent* if there is an orientation preserving homeomorphism $h: R^3 \rightarrow R^3$ such that $h(K_1) = K_2$. A link L is said to be *splittable* if there exists a smooth (or polyhedral) 3-ball B , an ordering of the components of the link K_1, K_2, \dots, K_m and an integer $0 < k < m$ such that $K_j \subset B$ for $j \leq k$ and $K_i \subset S^3 - B$ for $i > k$. A *diagram* for a knot or link K is an image of a regular projection (all self-intersections are non-tangential (transverse) and are double points) of K onto a plane with crossing information at each double point (p. 215, reference 2). Note that FIGURES 1 and 4 are examples of diagrams. Two knot or link diagrams D_1 and D_2 are said to be *equivalent* if D_1 can be obtained from D_2 by:

- 1) Deformations of the plane which do not alter the crossing information at each double point and
- 2) The three Reidemeister moves and their inverses. See FIGURE 2 for an illustration of these.

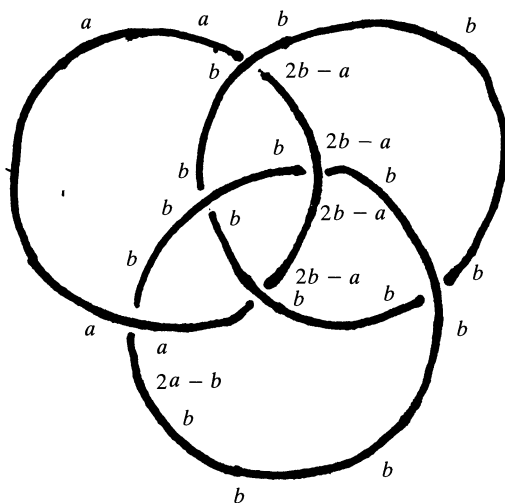


Figure 1

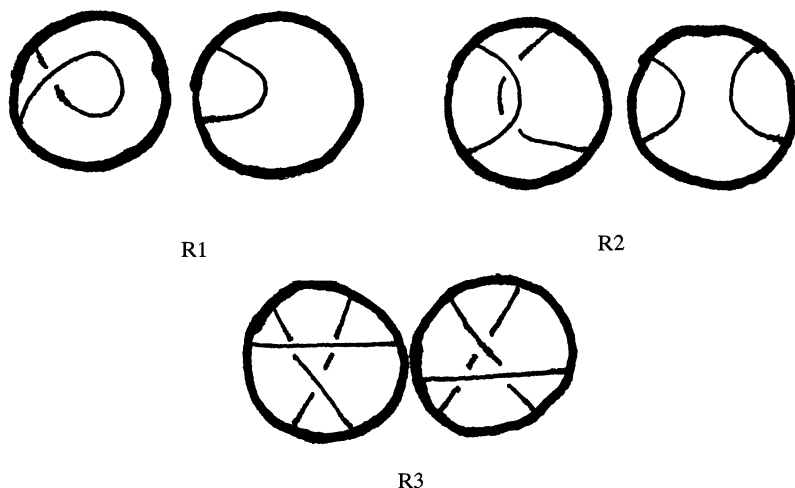


Figure 2. The Reidemeister Moves

2. THEOREMS. The following theorem is well known and will not be proved here.

Theorem 1. *Two knots or links are equivalent if and only if they have equivalent diagrams. See section 1B of reference [4] for a proof.*

A knot or a link K is said to be *colorable mod n* (n is assumed to be 3 or greater) if K has a diagram D in which it is possible to assign an integer to each arc of D which does not contain an undercrossing of D such that:

- 1) at each crossing we have $a + c = 2b \pmod{n}$ where b is the integer assigned to the overcrossing and a and c are the integers assigned to the other two arcs (see FIGURE 3) and
- 2) at least 2 distinct integers mod n are used in the diagram.

The following theorem is well known:

Theorem 2. *If K_1 is a knot or a link which is colorable mod n then every diagram of K_1 is colorable mod n .*

Proof: Exercise. All one has to check is: if a diagram D is colorable mod n and if one applies either a Reidemeister move (or its inverse) to D , the resulting diagram remains colorable mod n . \square

It follows from Theorem 1 and Theorem 2 that if K_1 is a knot or a link which is colorable mod n and K_2 is equivalent to K_1 , then K_2 is colorable mod n .

Corollary 3. *There exists a knot which is not equivalent to the unknot.*

Proof: Note that the trefoil knot (see FIGURE 4) is colorable mod 3 whereas the unknot is not. \square

We now come to the main result of this note:

Theorem 5. *If a link L is splittable then L is colorable mod 3.*

Proof: If L is splittable with a splitting ball B , then there exists a diagram for L in which the images of $L \cap B$ are separated from the images of $L \cap (S^3 - B)$ by a circle C . Give the components of the diagram of $L \cap B$ the monochrome coloring by assigning the integer 0 to each strand. Similarly, assign the strands of the diagram of $L \cap (S^3 - B)$ the integer 1. \square

It is an exercise to see that the standard diagram of the Borromean rings is not colorable mod n for any $n > 1$. The integer labeling of the diagram depicted in FIGURE 1 illustrates this: one has no choice but to set $a = b$. Thus we have an elementary proof that the Borromean rings link is unsplittable and thus the rings cannot be pulled apart.

Remark. If a knot or link K is colorable mod n , then one can obtain a homomorphism from $\pi_1(R^3 - K)$ onto the dihedral group $D_n = \{s, t | s^2 = 1 = t^n, sts = t^{n-1}\}$. This homomorphism is determined by the particular choice of coloring. See Kaufman [2] or Fox [5].

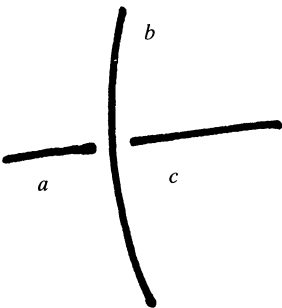


Figure 3

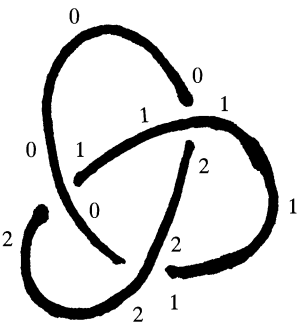


Figure 4

REFERENCES

1. B. Linström and H. Zetterström, Borromean circles are impossible, *American Math. Monthly*, 4(98) (1991) 340–341.
2. L. Kauffman, *Knots and Physics*, World Scientific, New Jersey, 1991.
3. M. A. Armstrong, *Basic Topology*, Springer-Verlag, New York, 1983.
4. H. Zieschang and G. Burde, *Knots*, de Gruyter, New York, 1985.
5. R. H. Fox, A Quick Trip Through Knot Theory, *Topology of Manifolds*, Prentice-Hall 1962.

Department of Mathematics
Bradley University
Peoria, IL 61625
onanyes@bradley.bradley.edu

Letter to the Editor:

Recently Grosof and Taiani [1] gave an algebraic proof that if $Q(X) = \prod_1^n (X - r_i)$ with the r_i distinct, then $\sum P(r_i)/Q'(r_i) = 0$ for $\deg(P) \leq n - 2$. I should like to add that this result has a home in algebraic number theory, as part of the computation of the “different”. The usual proof there [2, p. 135; 3, p.56; 4, p.144] is yet another ingenious algebraic argument. First, standard methods yield the partial fraction decomposition

$$1/Q(X) = \sum Q'(r_i)^{-1}/(X - r_i).$$

The right-hand side, as a formal power series in X^{-1} , is

$$\begin{aligned} X^{-1} \sum Q'(r_i)^{-1}/(1 - r_i X^{-1}) \\ = \sum [Q'(r_i)^{-1} r_i^k] (X^{-1})^{k+1}. \end{aligned}$$

But the left-hand side is

$$\begin{aligned} (X^n + a_1 X^{n-1} + \dots)^{-1} \\ = X^{-n} (1 + a_1 X^{-1} + \dots)^{-1} \\ = X^{-n} - a_1 X^{-(n+1)} + \dots \end{aligned}$$

Comparing terms, we recover the fact that $\sum r_i^k/Q'(r_i) = 0$ for $k \leq n - 2$; we also see that the sum is equal to 1 when $k = n - 1$.

References

1. M. S. Grosof and G. Taiani, Vandermonde strikes again. *Amer. Math. Monthly* 100 (1993), 575–7.
2. E. Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen*. Leipzig, 1923; reprint Chelsea, New York, 1970.
3. J. P. Serre, *Local Fields*. Springer-Verlag, New York, 1979. Translation of *Corps Locaux*, Hermann, Paris, 1962.
4. A. Weil, *Basic Number Theory*. Springer-Verlag, New York, 1967.

William C. Waterhouse
Department of Mathematics
The Pennsylvania State University
University Park, PA 16802

UNSOLVED PROBLEMS

Edited by: **Richard Guy and Richard Nowakowski**

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

Open Problems in Pattern Avoidance

James Currie

INTRODUCTION. What makes a mathematical area interesting? The area should contain a range of open problems: some very concrete and approachable, others “bigger”. These days it might help for the area to tie in with chaos and fractals. Finally, it couldn’t hurt for someone to offer cash for solutions to problems in the area.

A word w over an alphabet Σ is **nonrepetitive** (or squarefree) if no two adjacent blocks in w are identical. For example, the word $v = abcacb$ is nonrepetitive. On the other hand, the word $u = abc bcd$ is repetitive, since bc occurs next to itself in u . Early in this century the Norwegian number theorist Axel Thue showed that arbitrarily long nonrepetitive words can be formed using only three letters [25]. Since an infinite tree with finite branching must contain an infinite path, one can also find “infinite words” on three letters which are non-repetitive. We refer to these “infinite words” as ω -**words**.

Thue’s result has been rediscovered and republished a dozen times or more. One reason for this sequence of rediscoveries is that nonrepetitive sequences have been used to construct counterexamples in many areas of mathematics: ergodic theory, formal language theory, universal algebra and group theory, for example [16, 12, 6, 22].

WORDS AVOIDING PATTERNS. A **word** is a finite sequence of elements of some finite set Σ . We call the set Σ an **alphabet**, the elements of Σ **letters**. The set of all words over Σ is written Σ^* . We take a naive view of words as strings of symbols; thus the concatenation of two words w and v , written wv , is simply the string consisting of the letters of w followed by the letters of v . The **empty word**, with no letters, is denoted by ε .

Let S and T be alphabets. A **substitution** $h: S^* \rightarrow T^*$ is a function generated by its values on S . That is, suppose $w \in S^*$, $w = w_1 w_2 \dots w_n$ with $w_i \in S$, $i =$

$1, \dots, n$. Then $h(w) = h(w_1)h(w_2) \dots h(w_n)$. We do not allow $h(w_i) = \varepsilon$ for any i . As an example, we could give a substitution $h: \{1, 2, 3\}^* \rightarrow \{1, 2, 3\}^*$ by $h(1) = 123$, $h(2) = 13$, $h(3) = 2$. In this case, $h(123) = h(1)h(2)h(3) = 123132$.

A nonrepetitive word over Σ is said to **avoid** xx ; it cannot be written $ah(xx)b$ where $a, b \in \Sigma^*$ and $h: \{x\}^* \rightarrow \Sigma^*$ is a substitution. Thue also showed that arbitrarily long **cubefree** words on two letters exist [25]. Such words avoid xxx in the sense that they cannot be written $ah(xxx)b$. The infinite cubefree word discovered by Thue is referred to as the Morse-Thue sequence, and is an important example in *symbolic dynamics* [16, 21]. Symbolic dynamics is a key tool for studying chaos.

Before posing our problems, we need a bit more background. Let w and p be words. We say that w **contains** pattern p if we can write $w = ah(p)b$ for words a and b , and some substitution h . Otherwise, we say that w **avoids** p . Let a pattern p be fixed. Let Σ be an alphabet with k letters. If there are arbitrarily long words over Σ avoiding p , we say that p is **avoidable on** Σ . Clearly, only the number k of letters in Σ is significant here, so we also say that p is **avoidable on** k letters.

For example, xx is avoidable on 3 letters. We say that p is **unavoidable** if there is no k for which p is avoidable on k letters. For example, xyx is unavoidable. According to a pretty result of Zimin [26], a pattern p on n letters is avoidable if and only if Z_n avoids p , where Z_n is the word on $\{1, 2, \dots, n\}$ defined by $Z_1 = 1$, $Z_n = Z_{n-1}nZ_{n-1}$, $n > 1$. However, no method is known to determine the smallest alphabet on which p is avoidable [2, 3]. In [2], a word U_Δ is given which is avoidable on 4 letters, but not on 3. Perhaps all avoidable words are avoidable on 4 letters.

A word w is **strongly nonrepetitive** if no two adjacent blocks in w are permutations of each other. For example, $u = 512341231416$ is *not* strongly nonrepetitive since the adjacent blocks 12341 and 23141 are permutations of each other. Let p be a word over an alphabet Σ , $p = p_1p_2 \dots p_n$, $p_i \in \Sigma$, $i = 1, \dots, n$. Say that a word w **strongly avoids** p if we cannot write $w = a\hat{p}_1\hat{p}_2 \dots \hat{p}_nb$ where a, b are words, the \hat{p}_i are nonempty words, and \hat{p}_i is a permutation of \hat{p}_j whenever $p_i = p_j$. Thus a word is strongly nonrepetitive if and only if it strongly avoids xx .

It was known for some time that xx is strongly avoidable on 5 letters, but not on 3 letters [23]. It has recently been shown that xx is strongly avoidable on 4 letters [19]. On the other hand, the smallest alphabet on which xxx can be strongly avoided is the 3 letter alphabet and the smallest alphabet on which $xxxx$ can be strongly avoided is the 2 letter alphabet [10].

Let $a = a_1a_2a_3a_4 \dots$ and $b = b_1b_2b_3b_4 \dots$ be ω -words over some alphabet Σ , with $a_i, b_i \in \Sigma$. Define the distance between a and b to be $\rho(a, b) = (1/k)$ where $k = \min\{i \in \mathbb{N} | a_i \neq b_i\}$. Thus the longer a and b go on agreeing, the closer together a and b are. Let L be the set of nonrepetitive ω -words over $\Sigma = \{1, 2, 3\}$. With respect to the metric ρ , L has no isolated points; for any nonrepetitive ω -word a over Σ , we can find distinct nonrepetitive ω -words over Σ agreeing with a to as many places as desired [24]. It follows that L is a *Cantor set*.

Concrete Problems

1. Is there a pattern w which is avoidable on 5 letters but not on 4 letters? [2]
2. Let L be the set of nonrepetitive words over the 3 letter alphabet $\{1, 2, 3\}$. It is known that $c(n)$, the number of words of L of length n , grows exponentially [4]. Give an exact enumeration for L . For the solution to this problem I offer US\$100.

3. It is known [15] that the set of cubefree ω -words over a 2-letter alphabet is uncountable. Is the set a Cantor set?

“Bigger” Problems

1. Is there an algorithm which decides, given a pattern p and a natural number k , whether p is avoidable on k letters? [3] If so, give such an algorithm. I offer US\$100 for the solution to this problem.
2. Define **strongly avoidable** in the obvious way. Is there an algorithm which decides, given a pattern p , whether p is strongly avoidable? If so, give such an algorithm. Again, US\$100 to the solver of this problem.
3. Is there an algorithm which decides, given a pattern p and a natural number k , whether p is strongly avoidable on k letters? If so, give such an algorithm. I offer US\$100 for the solution to this problem.
4. For US\$100, decide the following conjecture: If pattern p is avoidable on Σ , then the set of ω -words on Σ avoiding p is a Cantor set.
5. For US\$100, decide the following conjecture: If the smallest alphabet on which p is avoidable is $\{1, 2, \dots, k\}$, then there exists a natural number m , and substitutions $f: \{1, 2, \dots, m\}^* \rightarrow \{1, 2, \dots, k\}^*$ and $g: \{1, 2, \dots, m\}^* \rightarrow \{1, 2, \dots, m\}^*$ such that $f(g^n(1))$ avoids p for every $n \in \mathbb{N}$.

REFERENCES

1. S. Arshon, Demonstration de l'existence des suites asymetriques infinies, *Mat. Sb.*, (N.S.) (2) 769–779; *Zbl.* 18, 115.
2. Kirby A. Baker, George F. McNulty and Wayne Taylor, Growth problems for avoidable words, *Theoret. Comput. Sci.* 69 (1989), no. 3, 319–345; *MR* 91f:68109.
3. Dwight R. Bean, Andrzej Ehrenfeucht and George McNulty, Avoidable Patterns in Strings of Symbols, *Pacific J. Math.* 85 (1979) 261–294; *MR* 81i:20075.
4. J. Brinkhuis, Non-repetitive sequences on three symbols, *Quart. J. Math. Oxford Ser. (2)* 34 (1983), 145–149; *MR* 84e:05008.
5. T. C. Brown, Is there a sequence on four symbols in which no two adjacent segments are permutations of each other? *Amer. Math. Monthly* 78 (1971), 886–888.
6. Stanley Burris and Evelyn Nelson, Embedding the dual of π_∞ in the lattice of equational classes of semigroups, *Algebra Universalis*, 1 (1971/72), 248–253; *MR* 45 #5257.
7. James D. Currie, Non-repetitive walks in graphs and digraphs, PhD thesis, University of Calgary (1987).
8. James D. Currie, Which graphs allow infinite non-repetitive walks? *Discrete Math.* 87 (1991) 249–260; *MR* 92a:05124.
9. James D. Currie, Subwords of non-repetitive words, *J. Combin. Theory Ser. A.*, to appear.
10. F. M. Dekking, Strongly non-repetitive sequences and progression-free sets, *J. Combin Theory Ser. A*, 27 (1979) 181–185; *MR* 81b:05027.
12. Andrzej Ehrenfeucht and Grzegorz Rozenburg, On the separating power of EOL systems, *RAIRO Inform. Théor* 17 (1983) 13–22; *MR* 84g:68059.
13. P. Erdős, Some unsolved problems, *Magyar Tud. Akad. Mat. Kutató. Int. Kozl.* 6 (1961), 221–254.
14. Roger C. Entringer, Douglas E. Jackson and J. A. Schatz, On non-repetitive sequences, *J. Combin. Theory Ser. A* 16 (1974), 159–164; *MR* 48 #10860.
15. Earl D. Fife, Binary sequences which contain no BBb , *Trans. Amer. Math. Soc.* 261 (1980), 115–136; *MR* 82a:05034.
16. Andrés del Junco, A transformation with simple spectrum which is not rank one, *Canad. J. Math.* 29 (1977) 655–663; *MR* 57 #6367.
17. Jacques Justin, Characterization of the repetitive commutative semigroups, *J. Algebra* 21 (1972), 87–90; *MR* 46 #277.
18. Juhani Karhumäki, On cube-free ω -words generated by binary morphisms. *Discrete Appl. Math.* 5 (1983), 279–297; *MR* 84j:03081.
19. Veikko Keränen, Abelian squares are avoidable on 4 letters, *Automata, Languages and Programming: Lecture notes in Computer Sciences* 623 (1992) Springer-Verlag, 41–52.

20. Filippo Mignosi, Infinite words with linear subword complexity, *Theoret. Comput. Sci.* 65 (1989), 221–242; *MR* 91b:68093.
21. Marston Morse and Gustav A. Hedlund, Symbolic dynamics I, II, *Amer. J. Math.* 60 (1938) 815–866; 62 (1940) 1–42; *MR* 1, 123d.
22. P. S. Novikov and S. I. Adjan, Infinite periodic groups I, II, III, *Izv. Akad. Nauk. SSSR Ser. Mat.* 32 (1968) 212–244; 251–524; 709–731; *MR* 39 #1531a–c.
23. P. A. B. Pleasants, Non-repetitive sequences, *Proc. Cambridge Philos. Soc.* 68 (1970) 267–274; *MR* 42 #85.
24. Robert O. Shelton and Raj P. Soni, Aperiodic words on three symbols I, II, III, *J. reine angew. Math.* 321; 327; 330 (1981) 195–209; 1–11; 44–52; *MR* 82m:05004a–c.
25. Axel Thue, Über unendliche Zeichenreihen, *Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiana* (1912) 1–67.
26. A. Zimin, Blocking sets of terms, *Mat. Sb. (N.S.)* 119 (161) (1982); *Math. USSR Sbornik* 47 (1984) 353–364.

Department of Mathematics
University of Winnipeg
Winnipeg, Manitoba
Canada R3B 2E9
currie@uwpg02.uwinnipeg.ca

Serendipity

After reading the letter to the editor from R. Norwood [The last math journal, *Amer. Math. Month.*, 1993, p. 491–2], I wonder what will happen to those of us who enjoy mathematics, do not have a computer and like to read on public transportation. How will one be able to browse through various items including The American Mathematical Monthly at one's leisure and above all come across the most interesting articles which are always next to those one had planned to read? Will serendipity end?

A. M. Herzberg
 Department of Mathematics
 and Statistics
 Queen's University
 Kingston, K7L 3N6
 CANADA

PROBLEMS AND SOLUTIONS

Edited by:

Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before March 31, 1994 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10330. *Proposed by R. Bruce Richter, Carleton University, Ottawa, Ontario, Canada, and Josef Širáň, Technical University of Bratislava, Bratislava, Slovakia.*

Let n and k be given positive integers. Define q, r, s, t to be the unique integers such that $n = qk + r = s(k + 1) + t$, with $0 \leq r < k$ and $0 \leq t \leq k$. Show that

$$\binom{q}{2}k + rq \geq \binom{s}{2}(k + 1) + ts.$$

10331. *Proposed by Carl Pomerance, University of Georgia, Athens, GA.*

Find all positive integers n such that $n!$ is multiply perfect; i.e., a divisor of the sum of its positive divisors.

10332. *Proposed by Kiran S. Kedlaya, student, Harvard University, Cambridge, MA.*

If n and k are integers with $0 \leq k \leq n$, prove that

$$\binom{2n}{n+k} = \sum_j 2^{n-k-2j} \binom{n}{j} \binom{n-j}{j+k}.$$

10333. *Proposed by Michael Golomb, Purdue University, West Lafayette, IN.*

For a positive integer n with $2^k \leq n < 2^{k+1}$, let $L(n) = 2^k$ ($k = 0, 1, 2, \dots$). Let $S(n)$ be the sum of the binary digits of n .

- (a) Evaluate $\sum_{n \geq 1} \frac{1}{L^2(n)S(n)}$.
- (b) Show that $\sum_{n \geq 1} \frac{1}{L(n)S(n)}$ diverges.
- (c) Show that $\sum_{n \geq 1} \frac{1}{L(n)S^{1+\delta}(n)}$ converges for every $\delta > 0$.

10334. *Proposed by John Sarli, California State University, San Bernardino, CA.*

Let M be a fixed n by n matrix with complex entries which is *not* nilpotent. For $a, b \in \mathbb{C}$, define the linear operator $M_{a,b}$ on the space of n by n complex matrices by $M_{a,b}(N) = aMN + bNM$. If the operators $M_{a,b}$ and $M_{c,d}$ have the same characteristic polynomial, show that $a^k + b^k = c^k + d^k$ for some k , $1 \leq k \leq n$.

10335. *Proposed by David Borwein, University of Western Ontario, London, Ontario, Canada, and Jonathan Borwein, Simon Fraser University, Burnaby, British Columbia, Canada.*

Let r be a positive constant and $c_0 \geq 0$. Consider the iteration

$$c_{n+1} = c_n + r - \frac{c_n}{\sqrt{1 + c_n^2}}.$$

- (a) For which values of r does the sequence $\langle c_n \rangle$ converge?
- (b) In case of convergence to c with $c \neq c_0$, prove that $\lim(c_{n+1} - c)/(c_n - c)$ exists and determine its value.
- (c) In case of divergence, find an asymptotic expression for c_n .

10336. *Proposed by Ignacy I. Kotlarski, Oklahoma State University, Stillwater, OK.*

Let X_1, X_2, \dots be a sequence of independent identically distributed random variables, each exponentially distributed with parameter a , $a > 0$, i.e., for $k = 1, 2, \dots$,

$$\mathbf{P}(X_k \leq x) = \begin{cases} 0 & \text{if } x \leq 0, \\ 1 - e^{-ax} & \text{if } x > 0. \end{cases}$$

Let B be a fixed Borel set in $[0, \infty)$ such that its Lebesgue measure $\mu_L(B)$ is finite and positive. Let

$$Y_k = X_1 + \dots + X_k$$

for $k = 1, 2, \dots$, and

$$\theta = \sum_{k=1}^{\infty} \mathbf{P}(Y_k \in B).$$

(a) Find θ as a function of a .

(b) Find a uniform minimum variance unbiased estimator of θ from a sample from the above exponential distribution of a fixed size n .

10337. *Proposed by Horst Alzer, Waldbröl, Germany.*

Let $n \geq 1$ be an integer. Let x_1, \dots, x_n be real numbers with $x_i \in (0, 1/2]$. Consider the statement

$$\prod_{i=1}^n \frac{x_i}{1-x_i} \leq \frac{\sum_{i=1}^n x_i^n}{\sum_{i=1}^n (1-x_i)^n}. \quad (\mathbf{F}_n)$$

(a) Prove \mathbf{F}_n for $n \leq 3$.

(b) Show that \mathbf{F}_n is false for $n \geq 6$.

(c) * What about \mathbf{F}_4 and \mathbf{F}_5 ?

NOTES

Notes: (10332) The sum may be considered as a sum over all integers j by using the convention that a binomial coefficient $\binom{a}{b}$ is zero unless $0 \leq b \leq a$. **(10337)** The inequality \mathbf{F}_n was suggested by the related statement

$$\prod (x_i/(1-x_i))^{1/n} \leq \sum x_i / \sum (1-x_i),$$

with $i = 1 \dots n$ in all sums and products. This statement is true for all $n \geq 1$ under the conditions given in the statement of the problem. More information on this inequality, due to Ky Fan, can be found in E. F. Beckenbach and R. Bellman, *Inequalities*.

SOLUTIONS

Alternating Parity in Chebyshev Systems

E 3456 [1991, 646]. *Proposed by A. S. Cavaretta, Kent State University, Kent, OH.*

Suppose $0 = m_0 < m_1 < \dots < m_n$ are integers such that $m_i \equiv i \pmod{2}$.

(i) Prove that a real polynomial

$$c_0 + c_1 x^{m_1} + \dots + c_n x^{m_n}, \quad \text{with } c_0 c_n \neq 0$$

has at most n real zeros, each zero being counted according to its multiplicity.

(ii) Prove that the generalized Vandermonde determinant

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_0^{m_1} & x_1^{m_1} & \cdots & x_n^{m_1} \\ \vdots & \vdots & \cdots & \vdots \\ x_0^{m_n} & x_1^{m_n} & \cdots & x_n^{m_n} \end{vmatrix}$$

is non-zero if x_0, x_1, \dots, x_n are any $n + 1$ distinct real numbers.

Solution by Thomas Kunkle, College of Charleston, Charleston, SC. Part (i): Let $p(x)$ be the polynomial in question. We say that the sequence

$$c_0, c_1, c_2, \dots, c_n \quad (1)$$

has a *sign change* at i ($0 \leq i < n$), if, for some $k \geq 1$, $c_i c_{i+k} < 0$, and if, for every j strictly between i and $i + k$, $c_j = 0$. By Descartes's Rule of Signs, the number of positive zeros of $p(x)$ (counted according to multiplicity) is at most the number of sign changes of (1), and the number of negative zeros is at most the number of sign changes of

$$c_0, -c_1, c_2, \dots, (-1)^n c_n. \quad (2)$$

Since $p(0) \neq 0$, to prove (i), we need only show that the number of sign changes of (1) and that of (2) sum to at most n .

By i we will always mean a nonnegative integer, strictly less than n , for which $c_i \neq 0$. By definition, a sign change can occur only at such i . Because $c_0 c_n \neq 0$, for every i there exists a $k = k(i)$ such that $c_i c_{i+k} \neq 0$, and, for all j strictly between i and $i + k$, $c_j = 0$. If $k = 1$, then exactly one of (1) and (2) will have a sign change at i , and if, instead, $k > 1$, then both (1) and (2) might have a sign change at i . Thus the total number of sign changes is less or equal to the number of i for which $k(i) = 1$ plus twice the number of i for which $k(i) \geq 2$. This is less or equal to $\sum_i k(i) = n$. This completes the proof of (i).

Part (ii): Suppose that the determinant is zero, or, equivalently, that there exists a nontrivial polynomial

$$p(x) = c_0 + c_1 x^{m_1} + \cdots + c_n x^{m_n}$$

vanishing at the points x_0, \dots, x_n . If c_0 is not zero, then, by part (i) of this problem, the number of real zeros of $p(x)$ cannot exceed $\max\{k: c_k \neq 0\} \leq n$, a contradiction. If, instead, c_0 is zero, we set $l := \min\{k: c_k \neq 0\}$, and rewrite $p(x)$ as x^{m_l} times

$$q(x) := c_l + c_{l+1} x^{m_{l+1} - m_l} + \cdots + c_n x^{m_n - m_l}.$$

Since at least n of the points x_0, \dots, x_n are nonzero, $q(x)$ has n real zeros. This also contradicts part (i), according to which $q(x)$ has at most $\max\{k - l: c_k \neq 0\} \leq n - 1$ real zeros. Thus the determinant cannot be zero.

Editorial comment. After the problem appeared, the proposer learned from E. Passow that part (i) had already appeared, with various generalizations, in E. Passow, "Alternating parity of Tchebycheff systems," *J. Approx. Theory* 9 (1973), 295–298. Related results are contained in E. Passow, "Extended Chebycheff systems on $(-\infty, \infty)$," *SIAM J. Math. Anal.* 5 (1974), 762–763. The solver suggested G. Pólya and G. Szegő, *Problems and Theorems in Analysis*, Vol. II, Springer-Verlag, 1972–76 for information on Descartes's Rule of Signs.

Solved also by D. W. Bailey, S.-J. Bang (Korea), D. Callan, R. J. Chapman (U.K.), P. Čížek (student, Czech Republic), T. C. Craven, M. Dindos (Slovakia), J. Duemmel, N. J. Fine, F. Flanigan, L. L. Gardner, H. W. Guggenheimer, Y. Ikeda, A. A. Jagers (The Netherlands), X. F. Jiang (China), I. Kastanas, K. S. Kedlaya (student), D. W. Koster, O. P. Lossers (The Netherlands), R. Martin (student), J. S. Muldowney (Canada), R. J. Neuhaus, J. H. Nieto (Venezuela), A. Nijenhuis, A. Pechtl (Germany), A. Pedersen (Denmark), F. C. Rembis, J. Rickert, M. Roth & O. Šuch (Canada), E. T. Wong, and the proposer.

Restricted Block-Walking

E 3465 [1991, 852]. *Proposed by Dragomir Ž. Đoković, University of Waterloo, Waterloo, Ontario, Canada.*

Let p , q , m , and n be given non-negative integers. Compute the number of sequences of $m + n + 1$ integers $k_{-m}, k_{-m+1}, \dots, k_{-1}, k_0, k_1, \dots, k_{n-1}, k_n$ satisfying

- (i) $-p \leq k_{-m} \leq k_{-m+1} \leq \dots \leq k_n \leq q$.
- (ii) $k_{-1} \leq 0 \leq k_1$.

Solution by William Y. C. Chen, Los Alamos National Laboratory, Los Alamos, NM. The answer is

$$\binom{m+p}{m} \binom{n+q+1}{n+1} + \binom{m+p}{m+1} \binom{n+q}{n}.$$

Recall that the number of nondecreasing sequences of r integers confined to an interval of s integers is $\binom{s+r-1}{r}$ (selections of r integers from s types with repetitions allowed). Now consider two cases: (1) $k_0 \geq 0$, or (2) $k_0 \leq -1$. In each case, the desired sequences are built by solving two selection problems. In Case (1), we have

$$-p \leq k_{-m} \leq \dots \leq k_{-1} \leq 0 \quad \text{and} \quad 0 \leq k_0 \leq k_1 \leq \dots \leq k_n \leq q.$$

In Case (2), we have

$$-p \leq k_{-m} \leq \dots \leq k_{-1} \leq k_0 \leq -1 \quad \text{and} \quad 0 \leq k_1 \leq \dots \leq k_n \leq q.$$

In Case (1), we take m elements from $p + 1$ and $n + 1$ from $q + 1$; in Case (2), we take $m + 1$ elements from $p + 1$ and n from $q + 1$. Together, we have the formula claimed.

Solved also by S.-J. Bang (Korea), J. C. Binz (Switzerland), D. Callan, R. J. Chapman (U.K.), M. Dindos (Slovakia), J. Fukuta (Japan), K. S. Kedlaya (student), E. F. Knapp, A. Nijenhuis, R. B. Richter (Canada), A. Tissier (France), M. Vowe (Switzerland), Anchorage Math Solutions Group, National Security Agency Problems Group, and the proposer. Two incorrect solutions were received.

Strong Fixed Points of Permutations

E 3467 [1991, 853]. *Proposed by Todd Feil, Denison University, Granville OH, and Gary Kennedy, Oberlin College, Oberlin OH.*

A permutation π on the set $\{1, 2, \dots, n\}$ is said to have j as a *strong fixed point* if $\pi(k) < j$ for $k < j$ and $\pi(k) > j$ for $k > j$. Let $h(n)$ be the number of permutations on $\{1, 2, \dots, n\}$ having at least one strong fixed point. Prove that

$$2(n-1)! - (n-2)! \leq h(n) \leq 2(n-1)!$$

for $n > 1$.

Solution by David Callan, University of Wisconsin, Whitewater, WI. For the lower bound, note that 1 and n are strong fixed points for $(n-1)!$ permutations, and $(n-2)!$ of these have been counted twice. For $n \leq 4$, equality holds, since 1 or n is a strong fixed point whenever 2 or $n-1$ is a strong fixed point. For $n \geq 5$, both inequalities are strict.

The permutations that do not fix 1 or n cannot strongly fix 2 or $n-1$. We bound the contributions for $3 \leq j \leq n-2$. The permutations that strongly fix j but not 1 or n permute $\{1, \dots, j-1\}$ without fixing 1 and $\{j+1, \dots, n\}$ without fixing n ; there are $[(j-1)! - (j-2)!][(n-j)! - (n-j-1)!] = (j-2)(j-2)!(n-j-1)(n-j-1)!$ of these. By comparing successive terms, one notes that this is maximized at the extremes. With $(n-4)$ choices for j , the additional contributions are bounded by $(n-4)(n-4)(n-4)! < (n-2)!$, as required.

Editorial comment. B. M. M. de Weger found the asymptotic expansion

$$2(n-1)! - (n-2)! + 2(n-3)! + 4(n-4)! + 22(n-5)! \\ + 125(n-6)! + 834(n-7)! + O((n-8)!)$$

for $h(n)$. Although there is no simple exact formula for $h(n)$, the generating function $\sum_{n=0}^{\infty} h(n)x^n = F(x)/(1 + xF(x))$, where $F(x) = \sum n!x^n$, appears in R. P. Stanley, *Enumerative Combinatorics*, Vol. I (Wadsworth and Brooks/Cole 1986), Exercise 32b, pages 49 and 61.

Solved also by R. J. Chapman (U.K.), W. Y. C. Chen, P. Čížek (student, France), R. High, N. Komanda, D. W. Koster, O. P. Lossers (The Netherlands), H. M. Marston, I. Praton, R. W. Sheets, A. Tissier (France), R. Tschiersch (Germany), D. B. Tyler, K. Wayland, B. M. M. de Weger (The Netherlands), National Security Agency Problems Group, University of Wyoming Problem Circle, and the proposer. Three incorrect solutions were received.

Primitive Trigonometric Power Sums

E 3468 [1991, 853]. *Proposed by Curtis Cooper, Central Missouri State University, Warrensburg, MO, Robert E. Kennedy, Central Missouri State University, and Stanley Rabinowitz, Westford, MA.*

Suppose m and n are positive integers such that all prime factors of n are larger than m .

(a) Prove that

$$\sum_{k \neq 1}^n \sin^{2m} \left(\frac{k\pi}{n} \right) = \frac{\phi(n) - \mu(n)}{4^m} \binom{2m}{m},$$

which $*$ denotes summation over integers relatively prime to n . (Here ϕ and μ denote the arithmetic functions of Euler and Möbius, respectively.)

(b) Find a similar formula for

$$\sum_{k=1}^n \cos^{2m} (k\pi/n).$$

Solution by Kevin Ford, student, University of Illinois, Urbana, IL. For part (b) we show that

$$\sum_{k=1}^n \cos^{2m} \left(\frac{k\pi}{n} \right) = \frac{\phi(n) - \mu(n)}{4^m} \binom{2m}{m} + \mu(n).$$

For part (a), the standard binomial expansion yields

$$\begin{aligned}
 \sum_{k=1}^n {}^* \sin^{2m} \left(\frac{k\pi}{n} \right) &= \sum_{k=1}^n {}^* \left(\frac{e^{\pi i k/n} - e^{-\pi i k/n}}{2i} \right)^{2m} \\
 &= \frac{(-1)^m}{4^m} \sum_{k=1}^n {}^* \sum_{j=-m}^m (-1)^{m+j} \binom{2m}{m+j} e^{(m+j-(m-j))ik\pi/n} \\
 &= \frac{1}{4^m} \sum_{j=-m}^m (-1)^j \binom{2m}{m+j} \sum_{k=1}^n {}^* e^{2\pi i j k/n}.
 \end{aligned}$$

If $j = 0$, then $\sum_{k=1}^n {}^* e^{2\pi i j k/n} = \sum_{k=1}^n {}^* 1 = \phi(n)$. If $j \neq 0$, the hypotheses of the problem imply that $(j, n) = 1$, hence as k runs through the set of reduced residues modulo n , so does $h = jk$. In this case,

$$\begin{aligned}
 \sum_{k=1}^n {}^* e^{2\pi i j k/n} &= \sum_{h=1}^n {}^* e^{2\pi i h/n} = \sum_{h=1}^n e^{2\pi i h/n} \sum_{d|(h,n)} \mu(d) \\
 &= \sum_{d|n} \mu(d) \sum_{l=1}^{n/d} e^{2\pi i l d/n} \\
 &= \mu(n) e^{2\pi i} + \sum_{\substack{d|n \\ d>1}} e^{2\pi i d/n} \left(\frac{1 - e^{2\pi i}}{1 - e^{2\pi i d/n}} \right) \\
 &= \mu(n).
 \end{aligned}$$

Splitting off the term for $j = 0$ first, we see that

$$\begin{aligned}
 \sum_{k=1}^n {}^* \sin^{2m} \left(\frac{k\pi}{n} \right) &= \frac{\phi(n)}{4^m} \binom{2m}{m} + \frac{\mu(n)}{4^m} \left((1 + (-1))^{2m} - \binom{2m}{m} \right) \\
 &= \frac{\phi(n) - \mu(n)}{4^m} \binom{2m}{m}.
 \end{aligned}$$

To obtain (b), we proceed as in part (a). This gives

$$\begin{aligned}
 \sum_{k=1}^n {}^* \cos^{2m} \left(\frac{k\pi}{n} \right) &= \sum_{k=1}^n {}^* \left(\frac{e^{\pi i k/n} + e^{-\pi i k/n}}{2} \right)^{2m} \\
 &= 4^{-m} \sum_{j=-m}^m \binom{2m}{m+j} \sum_{k=1}^n {}^* e^{2\pi i j k/n} \\
 &= \frac{\phi(n)}{4^m} \binom{2m}{m} + 4^{-m} \mu(n) \sum_{\substack{j=-m \\ j \neq 0}}^m \binom{2m}{m+j} \\
 &= \frac{\phi(n) - \mu(n)}{4^m} \binom{2m}{m} + \mu(n),
 \end{aligned}$$

since

$$\sum_{\substack{j=-m \\ j \neq 0}}^m \binom{2m}{m+j} = 2^{2m} - \binom{2m}{m}.$$

Editorial comment. The proposers included a reference to Stanley Rabinowitz, “Problem 1463,” *Crux Mathematicorum*, [1989, 207; 1990, 280], which dealt with a similar sum without the restriction to values of k relatively prime to n . In that form, one is essentially identifying the constant term of the Fourier series of $\cos^{2m}(x)$. This could be generalized to a use of the entire Fourier series of this function as a Discrete Fourier transform. Brian Conolly supplied a reference to B. W. Conolly and I. J. Good, “A table of discrete Fourier transform pairs”, *SIAM J. Appl. Math.*, 32 (1977), 810–822, which organizes work on this and similar formulas. In this context, the key steps in solving the present problem amount to the calculation of the transform of the characteristic function of a reduced set of residues modulo n .

Solved also by J. C. Binz (Switzerland), D. Callan, R. J. Chapman (U.K.), P. Čížek (student, France), C. Efthimiou, N. J. Fine, K. S. Kedlaya (student), D. W. Koster, L. E. Mattics, A. Pedersen (Denmark), G. Thompson, J. C. Vera Lizcano (Colombia), Anchorage Math Solutions Group, and the proposers.

An Identity Related to the Landen Transform

6672 [1991, 862]. *Proposed by H. B. Kushner, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY.*

If a and b are positive real numbers, prove that

$$\begin{aligned} \int_0^{\pi/2} \{(a \cos^2 \phi + b \sin^2 \phi)(a \sin^2 \phi + b \cos^2 \phi)\}^{-1/2} d\phi \\ = \int_0^{\pi/2} (a^2 \cos^2 \theta + b^2 \sin^2 \theta)^{-1/2} d\theta \end{aligned}$$

and use it to prove that the integral on the right is unchanged if a and b are replaced by $(ab)^{1/2}$ and $(a + b)/2$, respectively.

Solution by B. W. Conolly, Cambridge, U.K. Let $J(a, b)$ and $I(a, b)$ be the integrals on the left and right, respectively. The substitution $\tan \phi = \sqrt{b/a} \tan \theta$ shows that $J(a, b) = I(a, b)$. Moreover, the expression under the radical in $J(a, b)$ can be written

$$(a \cos^2 \phi + b \sin^2 \phi)(a \sin^2 \phi + b \cos^2 \phi) = a_1^2 \cos^2 2\phi + b_1^2 \sin^2 2\phi$$

where $a_1 = (ab)^{1/2}$ and $b_1 = (a + b)/2$. Thus

$$\begin{aligned} I(a, b) = J(a, b) &= \int_0^{\pi/2} (a_1^2 \cos^2 2\phi + b_1^2 \sin^2 2\phi)^{-1/2} d\phi \\ &= \frac{1}{2} \int_0^\pi (a_1^2 \cos^2 \theta_1 + b_1^2 \sin^2 \theta_1)^{-1/2} d\theta_1 \quad (\theta_1 = 2\phi) \\ &= \int_0^{\pi/2} (a_1^2 \cos^2 \theta_1 + b_1^2 \sin^2 \theta_1)^{-1/2} d\theta_1 = I(a_1, b_1). \end{aligned}$$

Editorial comment. Many solvers included material on the arithmetic-geometric mean or elliptic integrals in their proofs. To see the connection with elliptic integrals, assume $a < b$ and set $k = a/b$, $k' = \sqrt{1 - k^2}$. Then

$$\begin{aligned} I(a, b) &= \int_0^{\pi/2} (a^2 \cos^2 \theta + b^2 \sin^2 \theta)^{-1/2} d\theta = \frac{1}{b} \int_0^{\pi/2} (1 - k'^2 \cos^2 \theta)^{-1/2} d\theta \\ &= \frac{1}{b} \int_0^{\pi/2} (1 - k'^2 \sin^2 \theta)^{-1/2} d\theta = \frac{1}{b} K(k'), \end{aligned}$$

where $K = K(k)$ and $K' = K(k')$ are the usual complete elliptic integrals of the first kind for the modulus k . The *Landen Transformation* (see [1], [2] or [4]) states that

$$K' \left(\frac{2\sqrt{k}}{1+k} \right) = \frac{1+k}{2} K'(k). \quad (1)$$

With $k = a/b$, one can check that $(2\sqrt{k}/(1+k)) = a_1/b_1$. Many solvers observed that $I(a_1, b_1) = I(a, b)$ follows from the previous two equations.

Historically, elliptic integrals led to elliptic functions, which in turn led to elliptic curves. From the modern point of view, elliptic integrals are periods of an elliptic curve. To see how this works, consider the elliptic curve E defined by $w^2 = (a^2 + t^2)(b^2 + t^2)$. One can construct E by gluing together two copies of $\mathbb{C} \cup \{\infty\}$ which are cut from ia to ib and from $-ia$ to $-ib$. The cuts allow $w(t)$ to be given by a well-defined function on each sheet. A homology basis of E is $\{\gamma_1, \gamma_2\}$, where γ_1 is $\mathbb{R} \cup \{\infty\}$ on one copy of $\mathbb{C} \cup \{\infty\}$, and γ_2 consists of the segments from $-ia$ to ia on both copies of $\mathbb{C} \cup \{\infty\}$. Since

$$\frac{dt}{w} = \frac{dt}{\sqrt{(a^2 + t^2)(b^2 + t^2)}}$$

is a nonvanishing holomorphic form on E (unique up to a constant factor), the *periods* of E are the integrals

$$\int_{\gamma_1} \frac{dt}{w} = 2 \int_0^\infty \frac{dt}{w} \quad \int_{\gamma_2} \frac{dt}{w} = 4 \int_0^{ia} \frac{dt}{w}.$$

The substitution $t = b \tan \theta$ shows that $\int_0^\infty dt/w = I(a, b)$, so that $2I(a, b)$ is a period of the elliptic curve E . In terms of complete elliptic integrals of the first kind, the periods are

$$\int_{\gamma_1} \frac{dt}{w} = \frac{2}{b} K' \quad \int_{\gamma_2} \frac{dt}{w} = \frac{4i}{b} K.$$

We can also reconstruct E from its periods as follows. The quotient

$$\tau = \frac{\int_{\gamma_2} dt/w}{\int_{\gamma_1} dt/w} = \frac{2iK}{K'} \in \mathfrak{H} = \{x + iy \in \mathbb{C} : y > 0\}.$$

is sometimes called the *period* of E , and there is a complex analytic isomorphism $E \simeq \mathbb{C}/[1, \tau]$ which can be given explicitly in terms of the Jacobi elliptic functions sn , cn and dn . It follows that E is uniquely determined by τ modulo the action of $SL(2, \mathbb{Z})$ on \mathfrak{H} .

The identity $I(a, b) = I(a_1, b_1)$ relates E to the elliptic curve E_1 defined by the equation $w_1^2 = (a_1^2 + t_1^2)(b_1^2 + t_1^2)$. The substitutions used in the solution of the problem were $\tan \theta = \sqrt{a/b} \tan \phi$ and $\theta_1 = 2\phi$. To get to the elliptic curves, we use $t = b \tan \theta$ and $t_1 = b_1 \tan \theta_1$. The resulting change of variables is $t_1 = 2a_1 b_1 t / (a_1^2 - t^2)$, which comes from a map of the underlying elliptic curves since

$$t_1 = \frac{2a_1 b_1 t}{a_1^2 - t^2} \quad w_1 = \frac{a_1 b_1 w (a_1^2 + t^2)}{(a_1^2 - t^2)^2}$$

defines a function $\Phi: E \rightarrow E_1$. This map preserves the group structure of E and E_1 and is an example of what is called an *isogeny*.

The isogeny Φ is the key to the whole story. Since it has degree two, one period is preserved while the other is doubled. In fact, we have $2 dt/w = dt_1/w_1$, and since Φ maps a_1 to ∞ , and ∞ to 0, we obtain

$$\int_0^\infty \frac{dt}{w} = \int_0^{a_1} \frac{dt}{w} + \int_{a_1}^\infty \frac{dt}{w} = 2 \int_0^{a_1} \frac{dt}{w} = \int_0^\infty \frac{dt_1}{w_1}, \quad (2)$$

where one uses $t \mapsto ab/t$ to justify the second equality. This proves $I(a, b) = I(a_1, b_1)$, which in turn implies the Landen Transform (1). Furthermore, Φ takes ia to ia_1 , so that

$$2 \int_0^{ia} \frac{dt}{w} = \int_0^{ia_1} \frac{dt_1}{w_1}, \quad (3)$$

and thus the other period is doubled as claimed. One can check that this proves the other half of the Landen Transform, namely

$$K\left(\frac{2\sqrt{k}}{1+k}\right) = (1+k)K(k).$$

If we combine (2) and (3), we see that the period of E_1 is

$$\tau_1 = \frac{2 \int_0^{ia_1} dt_1/w_1}{4 \int_0^\infty dt_1/w_1} = 2 \frac{2 \int_0^{ia} dt/w}{4 \int_0^\infty dt/w} = 2\tau. \quad (4)$$

There is also a connection with the arithmetic-geometric mean of Gauss (see [1], [2] or [3]). We have $a_1 = \sqrt{ab}$, $b_1 = (a + b)/2$, and if we iterate this construction, we obtain

$$a_{n+1} = \sqrt{a_n b_n} \quad b_{n+1} = \frac{a_n + b_n}{2} \quad n = 1, 2, \dots$$

Since a and b are positive, these numbers converge to a common limit which is denoted $\mu = M(a, b)$ (one can also let a and b be arbitrary complex numbers, but convergence is more complicated in this case—see [3]). Then the identity $I(a, b) = I(a_1, b_1)$ implies

$$\begin{aligned} I(a, b) &= I(a_1, b_1) = I(a_2, b_2) = \dots = I(\mu, \mu) \\ &= \int_0^{\pi/2} \frac{d\theta}{\sqrt{\mu^2 \cos^2 \theta + \mu^2 \sin^2 \theta}} = \frac{\pi}{2\mu}, \end{aligned} \quad (5)$$

which proves that $I(a, b)M(a, b) = \pi/2$. Since there are other methods for proving this relation between $I(a, b)$ and $M(a, b)$ (see [2]), $I(a, b) = I(a_1, b_1)$ becomes a consequence of the obvious identity $M(a, b) = M(a_1, b_1)$. We can also study (5) from the point of view of the underlying elliptic curves. Let E_n be defined by $w^2 = (a_n^2 + t^2)(b_n^2 + t^2)$. Then (4) implies that the period τ_n of E_n is given by $\tau_n = 2^n \tau$, so that $\tau_n \rightarrow \infty$. This means that in the moduli space $\mathfrak{h}/SL(2, \mathbb{Z})$, the elliptic curves E_n are “converging” (the technical term is *degenerating*) to a rational curve. Thus the limit integral in (5) is an integral over a rational curve, which is why it is so easy.

The integrals $I(a, b)$ and $J(a, b)$ of this problem have other interpretations as periods. In particular, $I(a, b)$ is a period of a curve (of genus 1) with equations $u^2 = a^2 x^2 + b^2 y^2$ and $x^2 + y^2 = 1$, while $J(a, b)$ is a period of a curve (of genus 3) with equations $u^2 = (a^2 x^2 + b^2 y^2)(b^2 x^2 + a^2 y^2)$ and $x^2 + y^2 = 1$. The changes of variable in the integrals can then be explained in terms of functions between these curves.

1. G. Almkvist and B. Berndt, "Gauss, Landen, Ramanujan, the arithmetic-geometric mean, ellipses, π , and the *Ladies Diary*," this MONTHLY 95 (1988), 585–608.
2. J. Borwein and P. Borwein, *Pi and the AGM*, John Wiley & Sons, New York, 1987.
3. D. Cox, "The arithmetic-geometric mean of Gauss," *L'Enseign. Math.* 30 (1984), 275–330.
4. E. Whittaker and G. Watson, *A Course of Modern Analysis*, Fourth Edition, Cambridge University Press, Cambridge, 1963.

Solved also by J. Anglesio (France), F. Bachmann (Switzerland), S.-J. Bang (Korea), R. Betts (student), K. V. Bhagwat (India), P. Bracken (Canada), W. A. Businger (Switzerland), R. J. Chapman (U.K.), Y. Diao, M. Drešević & N. Cakić (Yugoslavia), Z. Guan & N. Passell, R. W. Hopper, D. Jespersen, I. Kastanas, P. Landweber, H. Lipman, N. J. Lord (U.K.), O. P. Lossers (The Netherlands), J. Melville (U.K.), G. Miller (student, Canada), A. Pechtl (Germany), C. E. Rieck Jr. & M. Q. Rieck, T. Schira (Germany), D. Trautman, R. L. Young, K. Zacharias (Germany), University of South Alabama Problem Group, and the proposer.

An All-Ones Problem

10197 [1992, 162]. *Proposed by Uri Peled, University of Illinois, Chicago, IL.*

Light bulbs L_1, L_2, \dots, L_n are controlled by switches S_1, S_2, \dots, S_n . Switch S_i changes the on/off status of light L_i and possibly the status of some other lights. Assume that if S_i changes the status of light L_j , then S_j changes the status of light L_i . Initially all the lights are off. Prove that it is possible to operate the switches in such a way that all the lights are on.

Solution by O. P. Lossers, Eindhoven University of Technology, Eindhoven, The Netherlands. Define the matrix A as

$$A_{ij} = \begin{cases} 1 & \text{if switch } S_i \text{ controls bulb } L_j \\ 0 & \text{otherwise.} \end{cases}$$

Then A is a symmetric $(0, 1)$ -matrix with all-one diagonal, and it should be proved that the all-one vector belongs to the column space of A , when calculated modulo 2.

More generally we shall prove that for a symmetric binary matrix A the diagonal \underline{d} belongs to the column space modulo 2, denoted $\text{Im } A$. In this form the problem occurs as "Problem 798," *Nieuw Archief voor Wiskunde*, (4) 9 (1991), 117–118. We give a different solution

$$\underline{d} \in \text{Im } A \text{ is equivalent to } (\text{Im } A)^\perp \subseteq \langle \underline{d} \rangle^\perp.$$

So let $\underline{x} \in (\text{Im } A)^\perp$, i.e. $\sum_{i=1}^n x_i A_{ij} = 0$ for all j .

Hence $\sum_{i=1}^n \sum_{j=1}^n x_i A_{ij} x_j = 0$, which, by symmetry of A reduces to $\sum_{i=1}^n x_i^2 A_{ii} = 0$. So $\sum_{i=1}^n x_i d_i = 0$ since $A_{ii} = d_i$ by definition and $x_i^2 = x_i$. Thus $\underline{x} \in \langle \underline{d} \rangle^\perp$ as required.

Editorial comment. Most solvers used a matrix interpretation as above, but a few worked directly with a graph with incidence matrix A . The proposer, in consultation with N. Alon and L. Lovász was able to trace this form of the result to an unpublished result of T. Gallai (see L. Lovász, *Combinatorial Problems and Exercises*, North-Holland, 1979, Exercise 5.17). Other readers provided references to K. Sutner, "The σ -game and cellular automata," this MONTHLY, 97 (1990), 24–34; F. Galvin, "Solution to problem 88-8," *Mathematical Intelligencer* 11

(1989), 31–32; and K. Sutner, “Linear cellular automata and the Garden-of-Eden,” *Mathematical Intelligencer* 11 (1989), 49–53 (especially theorem 3.2).

Solved by 41 readers and the proposer.

Products of Nilpotent Matrices

10200 [1992, 163]. *Proposed by Daniel Goffinet, St. Étienne, France.*

(a) Prove that a (square) matrix over a field F is singular if and only if it is a product of nilpotent matrices.

(b) If $F = \mathbb{C}$, prove that the number of nilpotent factors can be bounded independently of the size of the matrix.

Solution by Richard Stong, University of California, Los Angeles, CA. We will show that for any field F four nilpotent factors suffice. Clearly a product of nilpotent matrices is singular. Hence we need only decompose a singular matrix into nilpotent factors.

Lemma 1. *If A is a square matrix over F , then A is a product of two matrices that can put into Jordan canonical form with eigenvalues in F . If A is singular, we may further assume that the Jordan canonical forms have a final row and a final column of zeroes.*

Proof: Passing to a different basis, we may assume A breaks up into blocks of the form

$$B = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -b_0 & -b_1 & -b_2 & \cdots & -b_{r-1} \end{pmatrix},$$

where the characteristic polynomial $p(t) = t^r + b_{r-1}t^{r-1} + \cdots + b_0$ is a power of an irreducible polynomial. If $b_0 \neq 0$ (i.e., $P(t) \neq t^r$), consider the identity

$$B = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ c_1 & c_2 & \cdots & c_{r-1} & c_r \end{pmatrix}^{-1} \\ \times \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \\ -c_r b_0 & c_1 - c_r b_1 & \cdots & \cdots & c_{r-1} - c_r b_{r-1} \end{pmatrix}.$$

For any $\{c_i\}$ the first matrix can be put in Jordan canonical form since its characteristic polynomial is $(t - 1)^{r-1}(t - c_r)$. The characteristic polynomial of the second factor is $p'(t) = t^r + (c_r b_{r-1} - c_{r-1})t^{r-1} + (c_r b_{r-2} - c_{r-2})t^{r-2} + \cdots + c_r b_0$. By choosing the c_i appropriately we may assume this polynomial is

$(t-1)^r$, hence splits over F . If A is singular, then we also get some blocks of the form

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

For these use the identity

$$N = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ (-1)^r & (-1)^{r-1} & (-1)^{r-2} & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \\ \times \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Note that both sides have 0 as an eigenvalue of multiplicity one. Hence both of their Jordan canonical forms have a final row and final column of zeroes. This shows that A is the product of two matrices that can be put into Jordan canonical form and if A is singular, then both have a final row and final column of zeroes. (In fact, if the null space of A is r -dimensional, we get the final r rows and final r columns all zero. Another interesting observation is that if F is infinite the proof above can be modified to show that both factors are diagonalizable.) After a change of basis both these factors have the form

$$\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix},$$

where B is $(n-r) \times (n-r)$ upper triangular and 0 denotes a matrix of zeroes, $r \times r$, $(n-r) \times r$, or $r \times (n-r)$, as required. (In fact, B has only nonzero entries on and just above the diagonal.) The following lemma factors these.

Lemma 2. *Let A be an $n \times n$ matrix of the form*

$$\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix},$$

where B is an $(n-r) \times (n-r)$ upper triangular ($r \geq 1$). Then A is a product of two nilpotent matrices.

Proof:

$$A = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ I_{n-r} & 0 \end{pmatrix},$$

where I_{n-r} denotes the $(n-r) \times (n-r)$ identity matrix and 0 denotes a matrix

of zeroes, either $r \times r$, $(n - r) \times r$, or $r \times (n - r)$, as required. The first factor is strictly upper triangular the second strictly lower, hence both are nilpotent.

Applying these two lemmas solves the problem.

Editorial comment. Frank Schmidt and Pei Yuan Wu submitted references to Pei Yuan Wu, "Products of nilpotent matrices," *Linear Algebra Appl.* 96 (1987), 227–232. In this article, Wu proves that every singular complex matrix A is a product of two nilpotent matrices, except for the case where A is a 2 by 2 nonzero nilpotent matrix (in which case he shows that such an expression is *never* possible). Wu also provided a reference to T. J. Laffey, "Factorizations of integer matrices as products of idempotents and nilpotents," *Linear Algebra Appl.* 120 (1989), 81–93. In the introduction to this article, Laffey asserts that Wu's result could be extended to arbitrary fields. However, no solution giving a complete argument leading to fewer than four factors over a general field was received.

Solved also by I. Kastanas, J. Sangroniz (Spain), T. Zeanah (part b only), and the proposer.

Collaborating editors: *David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Jerrold R. Griggs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pfiefer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, and William E. Watkins.*

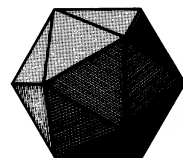
Answer to Picture Puzzle (p. 748)

No, they are not: they are Emile Borel and
Armand Borel.

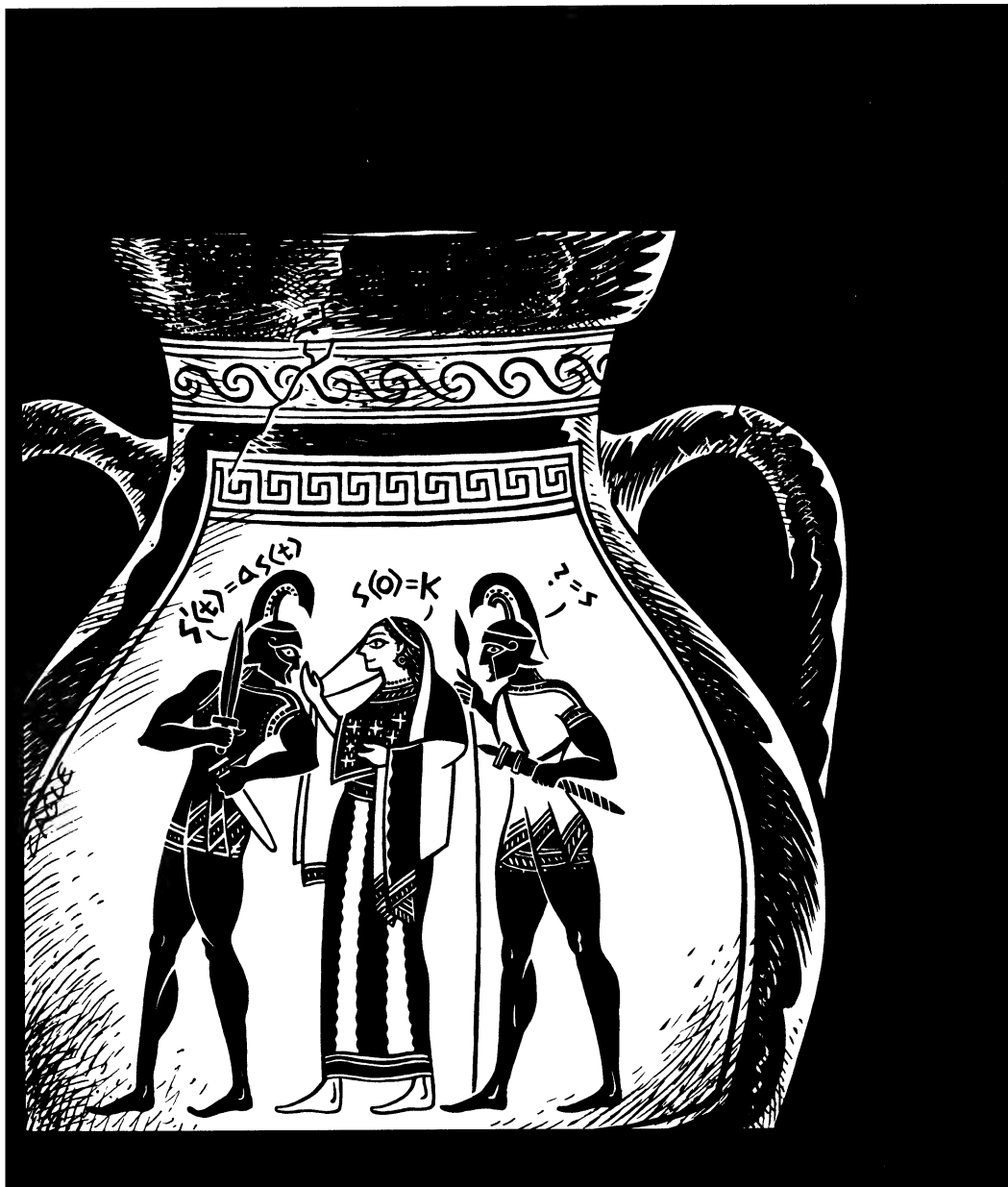
Dr. Marston Morse, professor of mathematics at Harvard University, has accepted a call to a professorship of mathematics at the Institute for Advanced Study at Princeton, New Jersey. The staff of the School of Mathematics now consists of the following members: Drs. Albert Einstein, Oswald Veblen, J. W. Alexander, John von Neumann, Herman Weyl and Marston Morse.

42(1935), 124

The American Mathematical Monthly



Volume 100, Number 9 / NOVEMBER 1993



NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING,
Department of Mathematics,
Indiana University,
Bloomington, IN 47405.

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International, Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1993, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

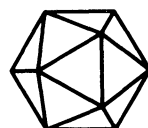
Cover:

The title is *ODE on a Grecian Urn*
by John Keats (1795–1821).

Blame and groans should be directed to
Tom Banchoff at Brown University.

The American Mathematical Monthly

Volume 100 Number 9 / NOVEMBER 1993
(ISSN 0002-9890)



Contents

ARTICLES

Thomas Archer Hirst—Mathematician Xtravagant V. London in the 1860s
/ J. HELEN GARDNER and ROBIN J. WILSON 827

From the Post-Markov Theorem Through Decision Problems to Public-Key
Cryptography / IRIS LEE ANSHEL and MICHAEL ANSHEL 835

Famous Nonmathematicians / STEVEN G. BUYSKE 845

The Fundamental Theorem of Linear Algebra / GILBERT STRANG 848

A Simple Proof of the Jordan-Alexander Complement Theorem /
ALBRECHT DOLD 856

Squaring the Circle with Holes / HANSKLAUS RUMMLER 858

FEATURES

COMMENTS 826

PICTURE PUZZLE 847

NOTES 861

COMPUTER SCIENCE SAMPLER

Parallel Addition / CATHERINE C. MCGEOCH 868

THE AUTHORS 872

PROBLEMS AND SOLUTIONS 875

REVIEWS

Second Year Calculus: From Celestial Mechanics to Special Relativity.

By David M. Bressoud / WILLIAM FARIS 884

TELEGRAPHIC REVIEWS 889

Thomas Archer Hirst— Mathematician Xtravagant V. London in the 1860s

J. Helen Gardner and Robin J. Wilson

After leaving the Academy I took my ticket for London by way of Dieppe and Newhaven . . . The passage was without exception the smoothest I ever made, the Channel was as quiescent as a duck-pond, the day beautiful and sunny . . . I was right glad to see the white cliffs of my native land and my eyes lingered gladly on the villages with their churches and on the farm-steads about which was an air of solid domestic comfort and prosperity which we look for in vain out of England. In short I felt a quiet pleasure in realising the fact that after long wanderings I was coming home at last and that sources of happiness were in store for me to which I had long been a stranger . . .

After two years establishing his reputation in Europe, Hirst decided that it was time to return home. On arrival in London, in the summer of 1859, he took up lodgings near John Tyndall.

9th October 1859: . . . Indeed my London life commences well. I have John close to me, can run into his rooms half an hour every evening and finish off the day with pleasant useful conversation. Never in my life was I better situated for getting through solid work and having the advantage of the best companionship. I trust the effects of all this will be visible by and bye . . .



James Joseph Sylvester (1814–1897)



Arthur Cayley (1821–1895)

By now, Tyndall was firmly established in the London scientific scene, and he introduced Hirst into his circle of friends. This enabled Hirst to become acquainted with the major scientific figures of the day. But what might have been merely polite introductions often developed further.

16th October 1859: On Monday having received a letter from [James Joseph] Sylvester I went to see him at the Athenaeum Club. We had an hour's talk in the little waiting room. He talked continuously for that time about his partitions of numbers and strange to say he was less obscure than I expected. He was, moreover, excessively friendly, wished we lived together, asked me to go live with him at Woolwich and so forth. In short he was excentrically affectionate . . .

Just before Christmas he called on Arthur Cayley, and spent a very interesting hour talking about Cayley's work on curves of the third order and a new method for obtaining the squares of the differences of the roots of a quintic, and his own work on derived curves of double curvature.

23rd December 1859: . . . I explained what I was doing in which he expressed some interest. I was a little amused and encouraged too by his asking me for a definition of the rectifying plane. The great geometer had forgotten it for the moment.

What a wonderful head he has, not merely round but spheroidal with the largest diameter parallel to his eyes, or rather to the line joining his ears. He never sits upright on his chair but with his posterior on the very edge he leans one elbow on the seat of the chair and throws the other arm over the back. Yet he is a keen sighted and extraordinary man, gentle I think by nature and at once timid, modest and reticent. Often when he speaks he shuts his eyes and talks as if he were reading from an unseen book, and talks well too so that one has to sharpen one's own wits to follow him.

His reading was wide, ranging from Alfred Tennyson's *Idylls of the King* to George Boole's *Differential Equations*. In the latter he found a passage that seemed to be related to his work.

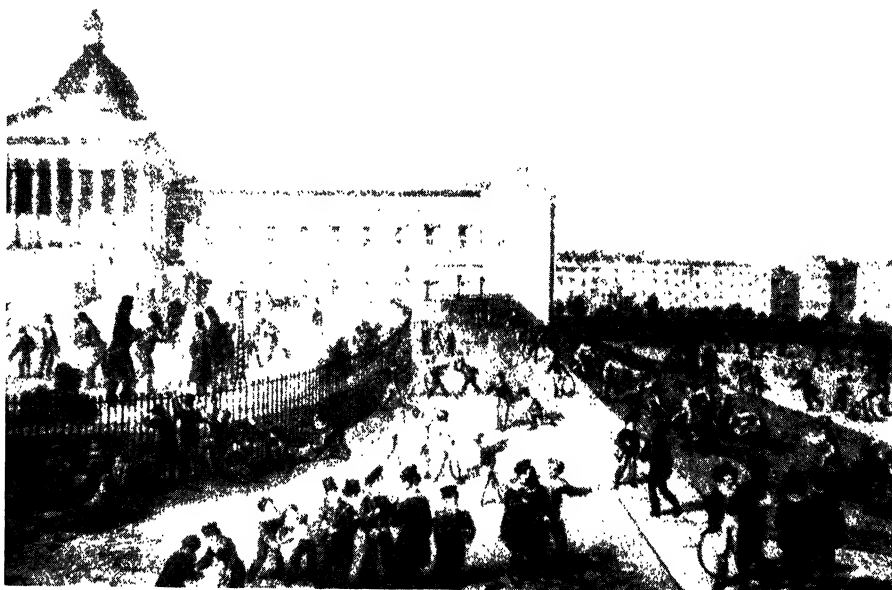
29th January 1860: . . . In the chapter on Partial Differential Equations p. 342 occurs this passage "Similar but more interesting applications may be drawn from the problem of the determination of equally attracting surfaces." This shows he has read my Memoir, but my name is not mentioned; yet I think I am the only one who has considered the problem in question.

Hirst also attended lectures of current importance, such as one given by his friend Thomas Huxley on Charles Darwin's recently proposed theory of evolution.

12th February 1860: . . . On Friday evening I heard Huxley's lecture on the Origin of Species at the Royal Institution. He gave us a noble peroration which is the part I shall remember longest . . . Tyndall introduced me to Babbage with whom we walked part of the way home.

Employment for a mathematician was as difficult to obtain as it had been seven years earlier when he returned from Berlin—but again, Hirst fell on his feet, being offered the post of mathematics teacher at the University College School. This was initially a temporary post, which Hirst accepted gladly. The headmaster was the distinguished classical scholar and mathematician Thomas Hewitt Key.

4th March 1860: . . . I was introduced by Key to my class on Wednesday morning at 9.15, and have continued to attend ever since. I am occupied there from 9.15 A.M. to 3 P.M. with an interval of an hour and a half at noon. My salary is £1 per day. For a school the instruction is of a superior kind. The highest class is engaged with the 6th book of Euclid, the Binomial Theorem in Algebra, De Moivre's theorem in Trigonometry and the simple machine in mechanics. So far I have succeeded quite well. I have merely been learning their powers.



University College School

Founded in 1833, this school was built on the site of University College, in Gower Street, London, where it remained until moving to its present location in Hampstead in 1907.

... I rise every morning now at 7, breakfast at 7.30, light my pipe at 8 and smoke and attend to other necessary matters connected with health until 8.30, then walk down to Regent Street where I take the Islington omnibus which puts me down at the end of Gower Street within a few minutes walk of the College. At noon I get a chop and glass of sherry where I can and return soon after 3.P.M. by the omnibus pretty well tired. Promising as my position is I should hesitate to accept it as a permanency. The consideration of £1 per day would not induce me to neglect my dear "derived surfaces".

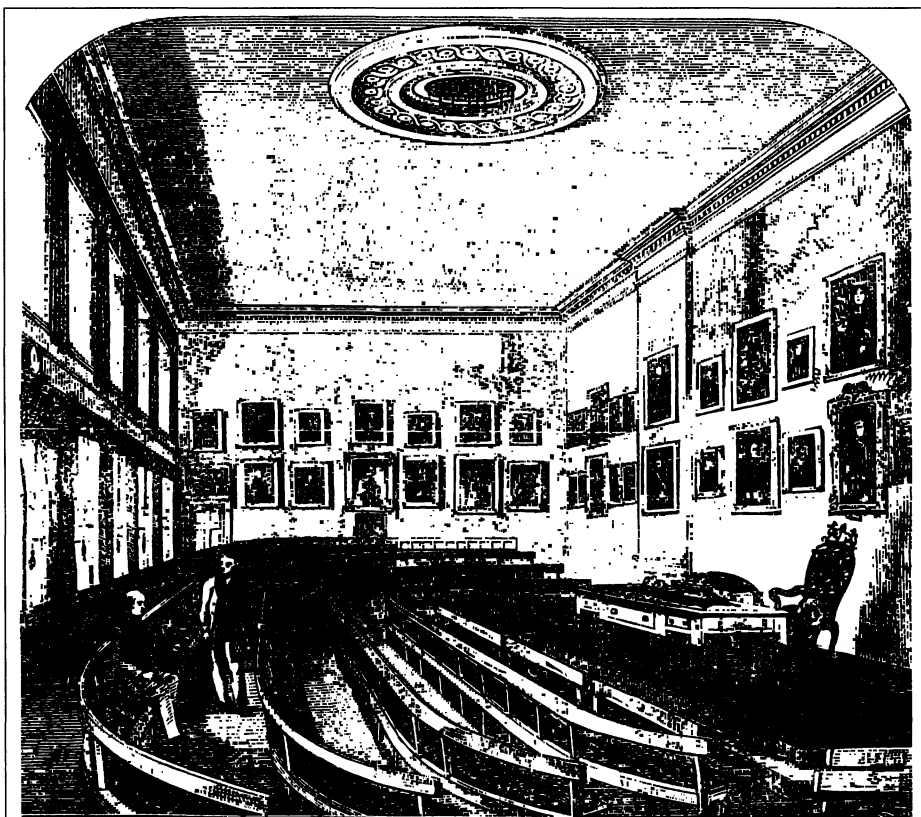
Not long after this, Hirst found himself drawn towards a rather different type of activity.

3rd June 1860: ... a week ago (Friday week) I became enrolled in the Volunteer Guards (6 feet men). I paid one guinea entrance and one guinea subscription in advance. I have been twice to drill once at the house of our Sergeant and Secretary Mr Halse who lives quite near, and once in undress uniform at Hungerford Hall. The uniform is exceedingly conspicuous, a red tunic with black belt and shoulder strap, a black patent-leather helmet (Prussian shape) with black plume and black trousers with red stripe. The undress is a red flannel jacket. I have joined them chiefly for the sake of the drill which I hope will be beneficial but I am also quite prepared to accept all the consequences and in case of need to defend my country with my life. As far as physique is concerned the Guards are a fine body of men about 70 in number, the uniform is expensive and consequently the members are all gentlemen.

In the meantime, Hirst continued, albeit slowly, with his translations and lecturing. His investigations had ground to a standstill, but even so, on March 1861, at the age of only 30, Hirst was nominated for a Fellowship of the Royal Society. His certificate was signed by (among others) Boole and Sylvester. There were many

other candidates, so he was not hopeful.

24th March 1861: ... the Royal Society were discussing my merits. I have two powerful competitors [Henry] Smith of Oxford and [James Clark] Maxwell of King's College; unless all three can be admitted I must expect to be the excluded one ... Yesterday I was at an evening party at Dr. Carpenter's and was introduced to Helmholtz and Maxwell with both of whom I had long conversations. The former is a little reserved, the latter talkative with a Scotch brogue, he took great interest in my ripples about which we spoke for some time ...



The Royal Society of London

The Royal Society was founded in 1662 by King Charles II. In 1863 it moved into these new rooms in Burlington House, Piccadilly. It is now located in Carlton House Terrace, near St. James's Park.

Figure 3.

Unfortunately, his health was causing him problems. Even after an Easter vacation, he felt exceedingly weak and spiritless. Although he suffered from no particular ailment, he had no energy for either his schoolwork or his researches.

14th April 1861: ... I pay the greatest attention to diet and avoid smoking all day. But instead of being better for such abstinence I feel weaker. I must persevere however for although the two pipes I still allow myself do me good at the time I have a firm belief that for my complaint, indigestion, smoking must be injurious or at least cannot be beneficial. Sooner or later therefore

abstinence must tell upon my health. Who knows how much of my present debility is due to the habit (of 11 or 14 years standing) of smoking five or six times a day. To cut down smoking to two pipes a day has been one of the greatest trials I have gone through, but perseverance diminishes the trial.

A week later, however, he learned from Tyndall that the Council of the Royal Society had placed his name upon the list of candidates to be elected Fellows in June. There were about 45 candidates, only 15 of which were chosen.

21st April 1861: ... Next morning I received the following note from Cayley:

Dear Sir,—I have much pleasure in being able to inform you of your name being on the Council list for the next election of Fellows of the Royal Society.

Believe me yours very sincerely
A. Cayley

... Of course the news was very welcome to me and in reply to Cayley I assured him that the honour would always be enhanced to me by the thought that *his* name was amongst those of the Council who had lent me so generous a support.

In March 1862 he recorded that he had been unable to get through his teaching work at the School. Extreme flatulence had produced giddiness which totally prevented him from standing at the board, and a strange numbness crept over his right arm and leg. It transpired that he had become very ill with dyspepsia, from which he was to suffer for over a month. Happily, he was soon back enjoying the company of his friends.

4th May 1862: ... Yesterday, Saturday, Cayley, Sylvester and Harley dined with me. Tyndall was not present. It was without question to me the most interesting dinner party I ever gave and I believe one of the most successful at least all appeared to enjoy themselves. I contrived to give my three guests opportunities of communicating their latest results. Cayley explained his late controversy with Boole on a question of Probabilities. Sylvester was eloquent on the subject of *Reseaux* which has now complete possession of him. According to his own confession he is so excited about it that he cannot trust his own critical judgment and has to call Smith of Oxford to his assistance. Harley entered into a few particulars on his Differential Resolvents of Algebraic Equations and I communicated my results on Derived Surfaces which Sylvester pronounced to be at once interesting and 'wonderful'. At 9.30 P.M. we all adjourned (in a cab) to Sabine's Soirée at Burlington House ...

He was now making good progress with his investigations. Shortly afterwards, he met Augustus De Morgan and told him of his researches, but received a very cool reception.

15th June 1862: ... He had no better remark to make than 'How did you come across that problem?' There are such an immense variety of similar questions. It was a kind of pooh pooh in fact. I felt angry with myself at having taken him even so much into my confidence. I ought to have *felt* that interest would not be reciprocal. A dry dogmatic pedant I fear is Mr. de Morgan notwithstanding his unquestioned ability ...

One of the most important scientific events of 1862 was the meeting of the British Association held at Cambridge, where he made several new acquaintances and presented a short communication on pedal curves which was 'listened to with attention but created no discussion'.

4th October 1862: ... I was much pleased with Boole ... Immediately after breakfast I stepped up to him and introduced myself. The same day we sat together at the Hall dinner and had some pleasant chat. Evidently an earnest able and at the same time a genial man.

He was, however, rather less impressed when he subsequently came across the distinguished physicist William Thomson, later Lord Kelvin.

7th June 1863: ... I have attended Thomson's two lectures at the Royal Institution on the Electric Telegraph. More random unsatisfactory lectures I never listened to.

15th June 1863: ... On Tuesday last I was at an "at home" given by Dr. and Mrs. King, the parents of one of my pupils and moreover relations of Prof. W. Thomson of Glasgow. It was the first time I had been introduced to Thomson. I cannot say that we suited one another very well or exchanged many words. He was civil and spoke flatteringly of my papers.

During the late summer, his health began to deteriorate again, making it difficult for him to transfer his thoughts to his researches, and he paid an extended visit to France, Switzerland, Germany, Italy, and Norway. His journal records his sadness at the death of Steiner, as well as meetings with both old and new acquaintances. While in Germany, he attended a gathering of the Naturforschende Gesellschaft, where he met Rudolf Clausius, whose memoirs he had earlier translated.

1st September 1863: ... I seized Clausius and he introduced me to Dedekind, a modest able mathematician Prof. at the Polytechnicum in Braunschweig. After dinner which was enlivened by numerous toasts, Clausius, Dedekind and I took our seats in a vehicle for the Excursion to the Morteratsch Glacier and a very pleasant excursion it was ...

The following summer, after completing his year's teaching, he set off on what was becoming an annual visit to the Continent. While in Paris, he dined with Chasles:

16th May 1864: ... The places of honour were given to Tchebichef whose acquaintance I renewed, for years ago I met him at Dirichlet's... On Wednesday Tchebichef called on me and left me some of his papers. He is evidently a good natured man, he has a stuttering way of speaking French and is lame.

... On Friday I took the American Railway to Sevres and sought the Maison Penel where Bertrand is at present residing. He had invited me to dinner... Tchebichef and myself were again the honoured guests on the right and left of Mrs Bertrand, Bertrand himself being opposite. I had much more conversation with Bertrand than I ever had before. I remember I had once a little prejudice against him. His manner I thought a little pretentious and forbidding. I begin to find that this is merely external, the man is kind at heart, extremely clever and full of *esprit*...

He particularly enjoyed spending a month in Bologna, where he renewed his acquaintance with Luigi Cremona and attended one of Cremona's lectures.

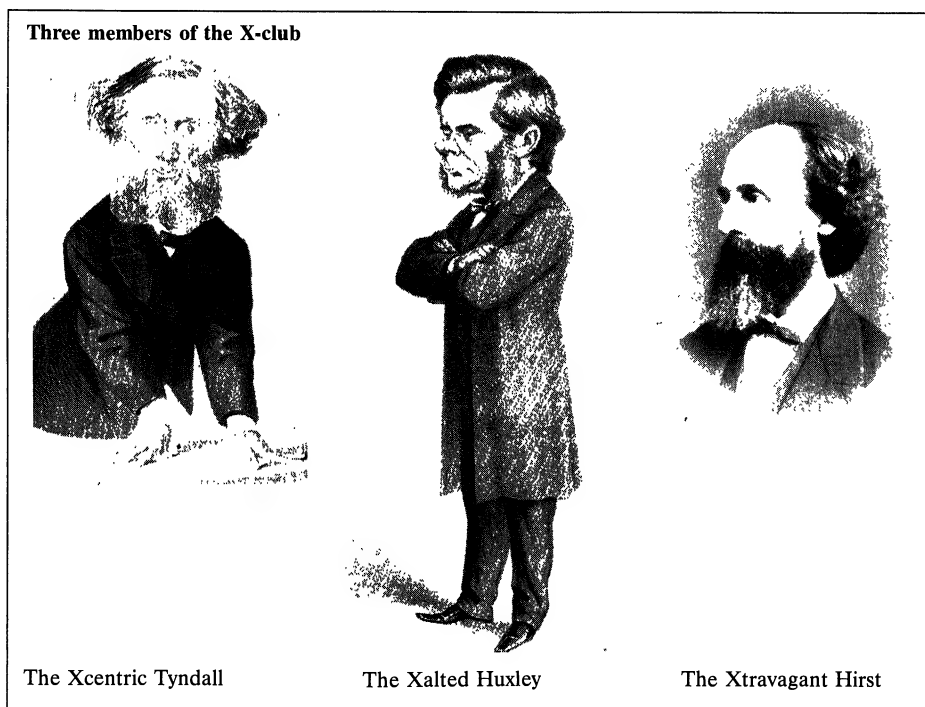
5th June 1864: ... He had a class of about 12 and lectured on the Theory of the Sun's Dial in connection with his Descriptive Geometry. He is evidently a good lecturer; everything was explained with perfect clearness. One peculiarity of the lecture arrangements was that instead of a black board on the side of the room the top of the table before the professor was of slate and on it he wrote and made figures in chalk. The figures were of course inverted to the audience.

In July 1864, he resigned his post at University College School in order to devote more time to research, and in November, he and Tyndall, along with other close friends, formed themselves into a select scientific club.

6th November 1864: ... On Thursday evening Nov. 3, an event, probably of some importance, occurred at the St George's Hotel, Albemarle Street. A new club was formed of eight members: viz: Tyndall, Hooker, Huxley, Busk, Frankland, Spencer, Lubbock and myself. Besides personal friendship, the bond that united us was devotion to science, pure and free, untrammelled by religious dogmas. Amongst ourselves there is perfect outspokenness, and no doubt opportunities

will arise when concerted action on our part may be of service. The first meeting was very pleasant and “jolly.” ... There is no knowing into what this club, which counts amongst its members some of the best workers of the day, may grow, and therefore I record its foundation. Huxley in his fun christened it the “Blasto dermic Club” and it may possibly retain the name.

The “jolly” time they had was obviously fruitful, for the X-club, as it became, was to influence the organization and image of English science for the next twenty years. They all acquired nicknames such as the Xcentric Tyndall, the Xalted Huxley and the Xtravagant Hirst.



In November 1864, Hirst was elected to the Council of the Royal Society for the first time. The very next week saw the first meeting of what was later to become the London Mathematical Society. Hirst became its first Vice-President, and was a member of the Council for almost two decades, becoming its treasurer, and later its President.

13th November 1864: ... On Monday last I attended the first meeting of the Mathematical Society at University College. De Morgan gave an address, which I seconded. I was put upon the Committee. I had at first declined but at De Morgan's request allowed my name to stand.

25th June 1865: ... On Monday I attended the Math. Soc. and proposed Cayley, Sylvester, Spottiswoode and Green as members. Sylvester gave us a capital communication on Newton's rule for the discovery of the imaginary roots of an equation.

On 18th August 1865, he received from the Secretary of University College the news of his appointment as Professor of Mathematical Physics. He had good reason to feel pleased with himself. His appointment to this newly-created chair established him as one of only seven physics professors in the country.

28th August 1865: ... Thus I have reached another step in my career. I have waited long for it and sacrificed much in order to stop in London. I trust I may have health and strength to perform my new duties efficiently.

15th October 1865: On Tuesday morning at 9 my work commenced with a lecture to 25 or 26 students. It passed off well and was listened to with the greatest interest. On Wednesday morning I commenced with my senior class, there were 5 students and a visitor... I have since continued my work every morning and have now altogether about 32 students which represent an income of 162 pounds upon which therefore I shall just be able to live without seeking for extra work.

This point marks the pinnacle of his career. As a Council member of the Royal Society, a member of the X-club, and Vice-President of the London Mathematical Society, he had become a most important member of the Scientific Establishment in London. Regrettably, his fortunes were soon to decline, as we shall see in the final article.

ACKNOWLEDGEMENTS. A typed version of the Thomas Hirst diaries is held at the Royal Institution in London, and quotations from the diaries appear here by courtesy of the Royal Institution. The diaries have been edited by W. H. Brock and R. M. MacLeod, and were published in microfiche by Mansell, London, in 1980.

The Open University
Milton Keynes MK7 6AA
England

Review

Infinitesimal Calculus. By F. S. Carey.
(Longmans Mathematical Series.) London,
Longmans, 1919. 8 vo. 20 + 352 + 9 pages.
Price 14 shillings.

The symbolism referred to for range and sequence is simple and worthy of mention. An open range from a to b is denoted by brackets $[a, b]$, a closed range by parentheses (a, b) ; and a range open only at one end by the appropriate combination of the bracket and parenthesis symbols; thus a range open at a and closed at b is denoted by $[a, b)$.

—*American Mathematical Monthly*
27, (1920) p. 470–471

From the Post-Markov Theorem Through Decision Problems to Public-Key Cryptography

Iris Lee Anshel and Michael Anshel

Dedicated to the Legacy of Emil Post

1. INTRODUCTION. On November 3–4 1988 a conference to commemorate the life and legacy of Emil Post (1897–1954), in anticipation of his one-hundredth birthday, was held at the City College of New York. Emil Post graduated from the City College in 1917 and received a Ph.D. from Columbia University in 1920. A postdoctoral fellowship at Princeton University was followed by long years of teaching in the public school system. He returned to the City College in 1935 as a member of its Mathematics Department where he resided for the remainder of his academic career. In the process Post was transformed from a brilliant young researcher into a great teacher and visionary intellectual. Four decades after their initial contacts with Post his former students spoke of him with a reverence that is rarely encountered in university life. The scholarly aspects of his commemorative meeting dealt with a wide range of Post's contribution to mathematics, logic and computer science. In this paper we should like to briefly recount his profound influence on the theory of algorithmic decision problems and the connections between this active field of research and current methods in public-key cryptography. We conclude our discussion by posing an historical question concerning the relationship between Post and the cryptologists of his day, the answer to which may shed new light on his legacy in the shadowy world of secret intelligence.

2. STRING REWRITING, THUE SYSTEMS, AND PRESENTATIONS. In this section we briefly review the basic concepts of string rewriting, Thue systems, and presentations. With this language in place we will be in a position to discuss some of the classical decision problems with which Post was concerned.

We begin by motivating this discussion with an historical example, a *Caesar cipher*. Identify the letters of the English alphabet $\{A, B, \dots, Z\}$ with the symbols $\{a_0, a_1, \dots, a_{25}\}$. Consider the set of pairs

$$\{(a_i, a_j) | i - 1 = j \bmod 26\}.$$

These pairs define a method of encrypting plaintext messages: for example 'IBM' ($a_8 a_1 a_{12}$) becomes 'HAL' ($a_7 a_0 a_{11}$) when a_i appearing in the plaintext string is replaced by a_{i-1} to obtain the ciphertext.

The idea behind the above example is to consider an alphabet together with a set of replacement or rewriting rules. With this in mind we begin our formal development. Let A denote a set of symbols (which we shall refer to as an *alphabet*). Consider $FM(A)$ the free monoid based on A . The elements of $FM(A)$ are the finite sequences of symbols or *words* from the alphabet A . Equality in

$FM(A)$ will be denoted by \equiv ; that is given words u and v in $FM(A)$, $u \equiv v$ if and only if they denote exactly the same string. Multiplication in $FM(A)$ of the words u, v is simply given by the concatenation uv , and the empty word e serves as the identity in $FM(A)$.

A *rewriting system* RW on A consists of the pair (A, P) where

$$P \subseteq FM(A) \times FM(A).$$

A *derivation* with respect to RW is a finite sequence of words in $FM(A)$,

$$w_1, \dots, w_n$$

such that either $n = 1$ or for each $i = 1, \dots, n - 1$ there exists

$$x_i, y_i \in FM(A)$$

and

$$(u_i, v_i) \in P$$

such that the equations

$$w_i \equiv x_i u_i y_i$$

and

$$w_{i+1} \equiv x_i v_i y_i$$

hold. We shall refer to the pair (u_i, v_i) as a *rewriting rule* or *production*, and term w_n *derivable* from w_1 .

Returning to the Caesar cipher rewriting system described above, the process of enciphering 'IBM' by 'HAL' is given by the following derivation:

$$a_8 a_1 a_{12}, a_7 a_1 a_{12}, a_7 a_0 a_{12}, a_7 a_0 a_{11}.$$

In his 1947 paper on the algorithmic unsolvability of a problem of Thue, Post introduced a class of combinatorial systems which he called *systems of semi-Thue type* [21]. From our perspective these are rewriting systems with finite alphabet and finitely many rewrite rules, together with a specified initial word. The semi-Thue systems serve as a convenient tool to represent the computation of a Turing machine. An exposition of this methodology is given in a now classic text on computability and unsolvability by Martin Davis [4], a student of Post and an authority on his life and work.

A *Thue system* T is a rewriting system such that the set P of productions is a symmetric relation on $FM(A)$: if $(u, v) \in P$ then $(v, u) \in P$. The imposition of this symmetry condition insures that the process of deriving (or *rewriting*) the word w_n from w_1 as above is reversible. Fixing our Thue system we now consider the equivalence relation P^* on $FM(A)$ generated by the set of productions P (by definition P^* is the intersection of those equivalence relations on $FM(A)$ which contain P). From the definition of P^* , it follows that two words $FM(A)$ are equivalent provided one is derivable from the other. Moreover P^* is a congruence on $FM(A)$. The semi-group $M(T)$ specified by the Thue system T is thus isomorphic to the factor monoid

$$M(T) \cong FM(A)/P^*.$$

Specifying each production (u, v) and its reverse by the equation $u = v$ we obtain the traditional monoid presentation for $M(T)$,

$$\langle A; u = v((u, v) \in P) \rangle. \quad (2.1)$$

Conversely, it is not difficult to show that given an arbitrary monoid M there exists

some Thue system T such that

$$M \cong M(T).$$

In the case A is finite we say that T and $M(T)$ are *finitely generated*, in the case P is finite we say that T and $M(T)$ are *finitely related*, and the Thue system T and the monoid $M(T)$ are said to be *finitely presented* if they are both finitely generated and finitely related. We will be concerned with finitely presented Thue systems T and we shall denote its associated monoid presentation by

$$\langle a_1, \dots, a_n; u_1 = v_1, \dots, u_k = v_k \rangle. \quad (2.2)$$

A *group alphabet* is an alphabet A partitioned into two disjoint subsets, $A(+)$, the *positive* symbols, $A(-)$, the *inverse* symbols together with an idempotent permutation inv of A such that

$$inv: A(+) \rightarrow A(-).$$

We write $a^{-1} = inv(a)$ for a in A and note $(a^{-1})^{-1} = a$. The notation extends to the words over the group alphabet by taking $e^{-1} = e$, and for $z = b_1 \dots b_i$ setting $z^{-1} = b_i^{-1} \dots b_1^{-1}$ where the b_i are positive or inverse symbols. A word z is said to be *freely reduced* provided it is not of the form $z = xaa^{-1}y$ or $xa^{-1}ay$ for any a in $A(+)$. A monoid presentation is said to be a *group presentation* provided it is specified by a Thue system over a group alphabet whose rewrite rules satisfy the following conditions:

- (i) All rewriting rules of the form $aa^{-1} = e$ and $a^{-1}a = e$ for a in A are contained in the system, and we call these rules *trivial relators*.
- (ii) All other rewriting rules or *non-trivial relators* occur in pairs of the form $u = e$, $u^{-1} = e$ where u, u^{-1} are free reduced words.

The collection of all rewriting rules is called the *defining relators*.

In general, the trivial defining relators are suppressed when specifying the group, as are the companions to each non-trivial relator $u = e$. In addition we list only the positive symbols. A finitely presented group G is thus specified and denoted by

$$\langle a_1, \dots, a_n; u_1 = e, \dots, u_k = e \rangle.$$

For an example of a presentation we consider F_n , the free group of rank n , which by definition has no non-trivial relators. We see that F_n is specified by the group presentation

$$\langle a_1, \dots, a_n; \rangle.$$

Every group is a factor group of a free group and may be specified up to isomorphism by a group presentation.

3. ALGORITHMIC DECISION PROBLEMS. By a *decision problem* we mean one whose instances require a yes/no answer. A decision problem is said to have an *algorithmic solution* if it is possible to program a digital computer to correctly supply the yes/no responses. If this is not possible we say that the problem is *algorithmically unsolvable*. If there is such a program and the running time of the program is bounded by a polynomial in the symbolic size of the input then we say the solution is *efficiently constructible*.

The algorithmic concepts referred to above are very natural to our computerized society. This was not the case in 1936 when Post [19] as well as Turing [24] both formulated a basis for these concepts by specifying idealized computing machines. In Turing's exposition a *universal* computer capable of executing any

algorithm is constructed. The *halting (or stopping) problem* for this machine is then shown to be algorithmically unsolvable. Post [21] clarified the construction of Turing's machine and applied his methods in order to resolve a problem of Thue [23] which we will discuss in §4.

Perhaps the most widely discussed decision problem of the twentieth century is Hilbert's Tenth Problem. This problem asks for an algorithmic solution for determining whether or not an integral polynomial equation has integer solutions. Post believed that Hilbert's Tenth Problem was algorithmically unsolvable [20]. The force of this belief was conveyed to Martin Davis who, along with Hilary Putnam and Julia Robinson, provided the basis for Yuri Matiyasevich's proof of its algorithmic unsolvability [14]. Martin Davis was cited by Marvin Minsky at his address to the City College conference as directing his (Minsky's) attention to Post's problem of *tag*. Minsky went on to show that this problem was algorithmically unsolvable [17]. Researchers such as Davis and Minsky have enabled Post's ideas to be transferred and transformed by succeeding generations. This process has made possible the enormous advances in *computer science* that have allowed it to emerge as an academic discipline.

4. THE POST-MARKOV THEOREM AND SUBSEQUENT DEVELOPMENTS.

We now restrict our attention to finitely presented monoids and groups. In the following discussion, we fix a presentation for the monoid or group in question. The *word problem* for a finitely presented monoid (resp. group) is to decide for arbitrary words, w, z in the alphabet (resp. group alphabet) whether or not the words are congruent via the associated Thue system (resp. group presentation).

Thue's word problem for finitely presented monoids appears in a 1914 paper of A. Thue [23] while the word problem for groups is formulated in the course of a topological investigation in 1911 by M. Dehn (see [3]). Another problem formulated by Dehn (see [3]) was the *conjugacy problem*: given arbitrary words w, z from a finite presentation of a group, decide if there is a word x such that w and $x^{-1}zx$ are congruent via the presentation (and thus define conjugate elements in the associated group). It is not difficult to prove that the algorithmic solvability of both the word and conjugacy problems is independent of the fixed finite presentation.

The negative resolution of the above problems represents an important achievement of twentieth century mathematics and one in which both Post and A. A. Markov played a fundamental role. In 1947 Post [21] and slightly later Markov [13] published independent proofs of the algorithmic unsolvability of the word problem for finitely presented Thue systems. The version of this result stated below reflects contemporary concern with constructive computational methods.

Post-Markov Theorem. *There exists finitely presented Thue systems having algorithmically unsolvable word problem. Moreover, there is an efficient algorithm P which, upon input of any Turing machine \mathcal{T} (resp. normal algorithm \mathcal{A}) will output a finitely presented Thue system $P(\mathcal{T})$ (resp. $P(\mathcal{A})$), such that $P(\mathcal{T})$ (resp. $P(\mathcal{A})$) has an algorithmically unsolvable word problem if \mathcal{T} (resp. \mathcal{A}) has an algorithmically unsolvable halting problem.*

A brief survey of finitely presented Thue systems with algorithmically unsolvable word problem is given by Matiyasevich [14] together with a proof of the Post-Markov Theorem employing normal algorithms. The rewriting techniques developed by Post to represent the computation of a Turing machine find their

way into several textbook proofs of the Novikov-Boone Theorem for the word problem in group theory.

Novikov-Boone Theorem. *There exists finitely presented groups having algorithmically unsolvable word problem. Moreover, there is an efficient algorithm B which upon input of any finitely presented Thue system T will output a finitely presented group $(B)T$ such that $B(T)$ has algorithmically unsolvable word problem if T has algorithmically unsolvable word problem.*

Examples of presentations of groups with algorithmically unsolvable word problem may be obtained using techniques studied by J. L. Britton (see [3]), but at the present time these presentations are quite complicated and involve many defining relators. The situation for finitely presented semigroups is however quite different. The semigroup

$$S = \langle a, b, c, d, e \mid ac = ca, ad = da, bc = cb, bd = db, eca = ce, edb = de, c^2a = c^2ae \rangle$$

while seeming simple in form has been shown by G. C. Tzeitlin to have an algorithmically unsolvable word problem (see Lallement [11] for an accessible proof). It is such striking examples that demonstrate the subtlety of the word problem.

The authors were surprised to discover a relationship between Thue's word problem and Dehn's conjugacy problem. A finitely presented *commutative* Thue system is one whose rewriting rules include (ab, ba) for all distinct a, b in its alphabet. Its associated semigroup is commutative and its word problem is algorithmically solvable. In M. Anshel [2] these properties are explicitly employed to show that the conjugacy problem for a special class of finitely presented groups is algorithmically solvable.

Results of a positive nature can be obtained for many large classes of groups and to give a perspective we highlight a few. A group G is termed *residually finite* provided when given $g \in G$, $g \neq 1$, there is a normal subgroup $N_g \triangleleft G$ such that g is not contained in N_g . Equivalently a group is residually finite if the intersection of the subgroups of finite index is the identity. The word problem for finitely presented residually finite groups is algorithmically solvable since both the words defining the identity element in G and the words defining the nonidentity elements in G are recursively enumerable (see [3]).

An important class of groups for which the word problem can be decided is the class of finitely generated groups with a single defining relator (e.g. one relator groups). Dehn originally formulated the word problem in the course of his investigation of the fundamental groups of orientable two dimensional manifolds (which are one relator groups). He did solve the word problem for these groups with the algorithm that has come to be known as *Dehn's algorithm*. Dehn's algorithm is studied geometrically in the context of *small cancellation* groups and more recently has surfaced in the study of *hyperbolic* groups initiated by Gromov (see [3]). The complexity of Dehn's algorithm is studied by B. Domanski and M. Anshel in [6] where it is shown that a finitely presented group of Dehn's algorithm has word problem solvable in linear time on a deterministic multitape Turing machine. W. Magnus (a student of Dehn) studied the entire class of one relator groups and proved through entirely algebraic means one of the landmark theorems in combinatorial group theory: the word problem for finitely generated one relator groups is algorithmically solvable (see [3]). More recently I. Anshel [1]

has investigated a class of groups with two relators and at least three generators and again the word problem is seen to be algorithmically solvable. The analysis here is close in spirit to that Magnus employed with the addition of methods from the theory of groups acting on graphs (see [3] for an introduction to these methods). To get some idea of the phenomenon that can arise when looking at this problem the reader is invited to consider the group E given by the presentation

$$\langle a, b | a^{-1}b^2a = b^3, b^{-1}a^2b = b^3 \rangle$$

and show every word in the generators defines the identity element (i.e. the group is trivial).

Recall that the conjugacy problem requires an algorithm to decide, given two elements in a group, whether or not they are conjugate. A striking result is proved by Charles F. Miller III in [16] regarding the conjugacy problem and is very much in the spirit of Post and of Miller's mentor W. W. Boone.

Miller's Theorem. *There exists finitely presented residually finite groups with algorithmically unsolvable conjugacy problem. Moreover, there is an efficient algorithm C which upon input of a finitely presented group G will output a finitely presented residually finite group $C(G)$, such that $C(G)$ has algorithmically unsolvable conjugacy problem if G has algorithmically unsolvable word problem.*

5. PUBLIC-KEY CRYPTOSYSTEMS BASED ON THE WORD AND CONJUGACY PROBLEMS. In conventional cryptography a method is provided to a sender S and a receiver R to transmit messages over an insecure channel by a mechanism such as a *code book* which provides easy encoding and decoding facilities. A particular weakness of such cryptosystems is that an interceptor with knowledge of the encoding facilities can readily decode transmitted messages. Conventional cryptography underwent automation at the end of World War I resulting in the development of mechanical cipher machines. These machines required the possession by both the sender S and the receiver R of a single key k in order to encode and decode a transmitted message. Cryptanalytic machine attacks (i.e. code-breaking) on one class of cipher machines, the Enigma, provided critical information to the Allies during World War II (see [8]–[10]).

One response to the advances in codebreaking technology was the introduction of *public-key cryptosystems*. This allows an R to receive messages from many senders S_1, S_2, \dots, S_n without the introduction of numerous codebooks or keys. These are replaced by a public-key mechanism which enables any sender S_i to easily encode messages which may be then transmitted over insecure channels. The code is designed so that if some third party T intercepts a message, T will find it computationally infeasible to break the code even with knowledge of the encoding mechanism employed by S_i unless the receiver's private key is known to T .

One widely discussed system is the RSA public-key cryptosystem named after its inventors R. L. Rivest, A. Shamir and L. Adelman (see [22]). Its security is generally thought to depend on the intractability of factoring large integers. One weakness, common to all public-key cryptosystems is that once the system is specified cryptanalytic attacks may be initiated. Two such attacks on the RSA system are outlined in [15].

Another more ambitious public-key cryptosystem based on the algorithmically unsolvability of the word problem was suggested by N. R. Wagner and M. R.

Magyarik (see [25]). Begin with a finitely presented group G specified by,

$$\langle a_1, \dots, a_n \mid u_1 = 1, u_2 = 1, \dots, u_m = 1 \rangle$$

with algorithmically unsolvable word problem, together with a *secret homomorphism*

$$h: G \rightarrow A$$

to a finitely presented group A with efficiently solvable word problem (such as a large finite group or finitely presented group of Dehn's algorithm). The homomorphism h is specified by its values on the finite generating set of G . Employing a consequence of Von Dyck's theorem [3], one can verify that h is a homomorphism by demonstrating that the image of each defining relator of G is the identity in the group A (note that this can be verified since A has efficiently solvable word problem). For this scheme we require two elements of G given respectively by words, y_0, y_1 such that

$$h(y_0) \neq h(y_1).$$

The mechanism for encoding is simply to replace each transmission of a '0' bit by any word y'_0 where

$$y'_0 = y_0 \bmod G$$

and similarly a '1' bit is replaced by any word y'_1 , such that $y'_1 = y_1 \bmod G$. Thus for example the sequence 0, 1, 1, 0 becomes

$$y'_0, y'_1, y''_1, y''_0$$

where y''_i is obtained from y'_i by successively inserting and/or deleting the defining relators of G . In this scheme the group G and the words y_0, y_1 constitute the public-key, while the homomorphism h constitutes the private-key. An interceptor T is faced with solving the word problem for G since R need never reveal the homomorphism $h: G \rightarrow A$. Although this system, like any other may be attacked, it is not based on such a fragile mechanism as the intractability of factorization within the current technology.

As a homage to Post we propose a public-key cryptosystem based on Miller's Theorem for constructing finitely presented residually finite groups with algorithmically unsolvable conjugacy problem (this extends the work of N. R. Wagner and M. R. Magyarik). The proposed cryptosystem begins with a finitely presented residually finite group G specified by,

$$\langle a_1, \dots, a_n; u_1 = e, \dots, u_k = e \rangle \quad (5.1)$$

with algorithmically unsolvable conjugacy problem (such a group's existence is insured by Miller's theorem, see §4). The additional data required for this system are two elements of $G, \{w, z\}$ such that

$$w \neq 1$$

and

$$z = 1$$

in G . Since G was chosen to be residually finite there exists a finite image of $G, G/N_w$ such that

$$w \notin N_w.$$

Thus when we consider the homomorphism

$$h: G \rightarrow G/N_w$$

we may assert that

$$\begin{aligned}h(w) &\neq 1 \\h(z) &= 1.\end{aligned}$$

Hence we conclude that w, z and $h(w), h(z)$ are non-conjugate pairs in G and G/N_w , respectively. We keep the homomorphism h secret and assume that computation in the finite group is efficient enough to determine when two elements specified by words are conjugate or whether a word defines the identity in G/N_w . The mechanism for recoding now follows that of the word problem cryptosystem described above with the enhancement of conjugation by words in G (as well as insertions and deletions of defining relators) being allowed. We observe that the complexity of the word problem for finitely presented residually finite groups is unknown at the time of this writing. In the Kourouva Notebook ([12] p. 58), F. B. Cannonito asks:

Do there exist finitely presented residually finite groups with recursive, but not primitive recursive, solution of the word problem?

As with the RSA cryptosystem, the conjugacy problem cryptosystem is based on the computational complexity of a special problem. To date, the research on integer factorization is massive [18] as compared to the research on the word problem for residually finite groups. In fact very little is known regarding the computational complexity of the word problem for these groups as the above recursion-theoretic problem indicates.

6. POST'S RELATION TO THE CRYPTOLOGY AND CRYPTOLOGISTS OF HIS ERA. We conclude our discussion by posing the following historical question:

What impact did Post have on the cryptologists of his era?

This question arises from two distinct sources. The first source is the very strong connection between the development of both the theory and practice of digital computation and cryptology and the second concerns Post's contemporaries at the City College of New York, an institution where Post spent nearly his entire adult life.

The intertwining of computation and cryptology is quite explicit in the lives of two individuals, Charles Babbage (1791–1871) and Alan M. Turing (1912–1954). Babbage was a prominent British mathematician whose Difference Engines and Analytic Engines were forerunners of the modern digital computer. He was also prominent among the cryptologists of his era for successful cryptoanalytic attacks on polyalphabetic ciphers (see [7]). Turing, a British contemporary of Post, played an instrumental role in the Allied victory in World War II by employing computing machines to break the Enigma code. Turing's life and work are documented in [9]. Post was certainly aware and indeed employed Turing's work with regard to the algorithmic unsolvability of the halting problem. It is only natural to ask whether there was a reciprocal interest in Post's work on the part of Turing (from the perspective of cryptology).

It is pointed out in [7] that a contemporary of Post, Charles J. Mendelsohn and faculty member of the History Department at the City College was very much involved in cryptological pursuits. In 1918 Mendelsohn was made a Captain in the Military Intelligence Division of the General Staff of the U.S. Army in charge of

decipherment of German codes. Mendelsohn was a classics scholar as well as historian and he pursued a lifelong study of historical ciphers and their originators including a study of Vigenère which appeared in 1940, the year following his death. In fact the proofs of this paper were corrected by his associate and friend, Lt. Col. William F. Friedman, the Principal Cryptoanalyst in the Office of the Chief Signals Officer of the U.S. Army. It was the same Friedman who rebuilt the U.S. cryptanalytic capability during the 1930's by hiring for the Signals Intelligence Service, Abraham Sinkov and Solomon Kullbeck (both City College graduates and both to go on to doctorates in mathematics and productive cryptological research for the National Security Agency). Friedman is regarded as one of America's top codebreakers in that his work led to the Japanese defeat at Midway during WWII (see [10]).

After discussions with the historian of cryptology David Kahn, and Harold Highland, editor emeritus of the journal *Computers and Security* (who attended City College during the nineteen thirties) there is a clear sense that the informal discussion groups which took place during that period would have lent themselves to consideration of Post's work. Further indications of such contact were evident when, in the course of his address to the November 1988 conference at City College, Marvin Minsky observed that Post rewriting methodology had been employed during the nineteen sixties on a cryptographic project.

The shadowy world of secret intelligence has provided scant information for an historical investigation of these matters. Even such a distinguished mathematician as Peter Hilton reports in [8]:

"I am unfortunately obliged to be reticent about the details of the work we did at Bletchley Park in breaking the highgrade German cipher (sic during WWII). For reasons best known—indeed, almost exclusively known—to themselves, the bureaucrats in Washington and Whitehall steadfastly refuse to declassify such details."

Steven Brams, the noted game theorist and political scientist, has remarked to us that the life and legacy of Emil Post represents one aspect of New York intellectual life during the first half of the twentieth century that is very much in need of deeper exploration. The authors hope that this paper serves to further this pursuit.

The authors would like to thank Martin Davis for supplying us with a preliminary manuscript [5] of his biographical and scholarly survey of Post's life and achievements.

REFERENCES

1. I. Anshel, A Freiheitssatz for a class of two-relator groups, *Journal of Pure and Applied Algebra* 72 (1991), 207–250.
2. M. Anshel, The conjugacy problem for HNN groups and the word problem for commutative semigroups, *Proc. of the Amer. Math. Soc.*, 61 no. 2 (1976), 223–224.
3. D. E. Cohen, *Combinatorial Group Theory: A Topological Approach*, Cambridge University Press, New York (1989).
4. M. Davis, *Computability and Unsolvability*, Dover Publications, Inc., New York (1982).
5. M. Davis, Emil Post: His Life and Work, manuscript.
6. B. Domanski and M. Anshel, The complexity of Dehn's algorithm for word problems in groups, *J. Algorithms* 6 (1985), 543–549.
7. O. I. Franksen, *Mr. Babbage's Secret: The Tale of a Cypher- and APL*, Prentice Hall Inc., Englewood Cliffs, New Jersey (1985).

8. P. Hilton, Working with Alan Turing, *The Mathematical Intelligencer* 13, no. 4 (1991), 22–25.
9. A. Hodges, *Alan Turing: The Enigma*, Simon and Schuster, New York (1983).
10. D. Kahn, *The Codebreakers: The Story of Secret Writing*, Macmillan, New York (1967).
11. G. Lallement, *Semigroups and Combinatorial Applications*, John Wiley and Sons, New York (1979).
12. L. J. Leifman and D. J. Johnson trans. ed., *The Kurovka Notebook, Unsolved Problems in Group Theory*, AMS trans., series 2, vol. 121 (1983).
13. A. A. Markov, On the impossibility of certain algorithms in the theory of associative systems, I.C.R. (Dokl.) *Acad. Sci. URSS* 55 (1947), 583–586 (English).
14. Yu. V. Matiyasevich, On investigations of some algorithmic problems in algebra and number theory, *Proc. Steklov Inst. Math.*, 168, issue 3, Amer. Math. Soc. Translation (1986), 227–252.
15. K. S. McCurley, Odds and ends from cryptology and computational number theory, in *Cryptology and Computational Number Theory*, C. Pomerance ed., *Proceedings of the Symposia in Applied Math.*, 42, AMS, Providence R.I. (1990), 145–166.
16. C. F. Miller III, *On Group-Theoretic Decision Problems and Their Classification*, Annals Math. Studies, Number 68, Princeton University Press, Princeton, New Jersey (1971).
17. M. L. Minsky, Recursive unsolvability of Post's problem of “tag”, and other topics in the theory of Turing machines, *Annals of Math.* 74 (1961), 437–453.
18. C. Pomerance, Factoring, in *Cryptology and Computational Number Theory*, C. Pomerance ed., *Proceedings of the Symposia in Applied Math.*, 42 AMS, Providence, R.I. (1990), 27–47.
19. E. L. Post, Finite combinatory processes—formulation 1, *J. Symbolic Logic* 1, (1936), 103–105.
20. E. L. Post, Recursively enumerable sets of positive integers and their decision problems, *Bull Amer. Math. Soc.* 50 (1944), 284–316.
21. E. L. Post, Recursive unsolvability of a problem of Thue, *J. Symbolic Logic*, 12 (1947), 1–11.
22. R. L. Rivest, A. Shamir and L. Adelman, A method for obtaining digital signatures and public-key cryptosystems, *Comm. of the ACM* (1978), 120–126.
23. A. Thue, Probleme Über Veränderungen von Zeichenreihen nach gegebenen Regeln, *K.V.S.S. No. 10*, (1914).
24. A. M. Turing, On computable numbers, with an application to the Entscheidungsproblem, *Proc. London Math Soc.* (2) 42 (1936–37), 230–265, A Correction 43 (1937), 544–546.
25. N. R. Wagner and M. R. Magyarik, A public key cryptosystem based on the word problem, in *Advances in Cryptology: Proceedings of Crypto 84*, ed G. R. Blakely and D. Chaum, Lecture Notes in Computer Science, 196 Springer Verlag, New York (1985), 19–36.

Department of Mathematics
Columbia University
New York, NY 10027

Department of Computer Science
The City College, CUNY
New York, NY 10031

There was a professor of Trinity
 Who found the square root of infinity;
 But in counting the digits
 He was seized with the fidgets,
 Dropped Science and took to Divinity.

—*American Mathematical Monthly*
 28, (1921) p. 394

Famous Nonmathematicians

Steven G. Buyske*

We often tell our students that there are many things besides teaching and actuarial work that they can do with a degree in mathematics, but I don't think they believe us. Over the years I've put together a list of well-known people who were math majors (or some equivalent in other countries and times), although not all of them completed their degrees. It's the most popular thing I've ever had on my office door. When I began this list, it had mostly contemporary Americans, and I called it "People who majored in math." Some of my students added their own names to their copies and posted them on their dorm doors.

I'd be delighted to hear of any additional names.

THE PUBLIC REALM.

Ralph Abernathy, civil rights leader and Martin Luther King's closest aide.

Corazon Aquino, former President of the Philippines. She was a math minor.

Harry Blackmun, Associate Justice of the US Supreme Court, AB *summa cum laude* in mathematics at Harvard.

David Dinkins, Mayor of New York, BA in mathematics from Howard.

Alberto Fujimori, President of Peru, MS in mathematics from the University of Wisconsin-Milwaukee.

Ira Glasser, Executive Director of the American Civil Liberties Union, both a BS and an MA.

Lee Hsien Loong, Deputy Prime Minister of Singapore, a Bachelor's from Cambridge.

Florence Nightingale, pioneer in professional nursing. She was the first person in the English-speaking world to apply statistics to public health. She was also a pioneer in the graphic representation of statistics; the pie-chart was her invention, for example. Not really a math major, she was privately educated, but pursued mathematics far beyond contemporary standards for women.

Paul Painlevé, President of France in the early 20th century, and one of the first passengers of the Wright Brothers. A ringer: he had a distinguished mathematical career.

Carl T. Rowan, columnist for the *Washington Post*.

Laurence H. Tribe, Professor at Harvard Law School, often regarded as one of the great contemporary authorities on Constitutional Law. An AB *summa cum laude* in mathematics from Harvard.

Leon Trotsky, revolutionary. He began to study Pure mathematics at Odessa in 1897, but imprisonment and exile in Siberia seem to have ended his mathematical efforts.

*I'd like to thank my colleagues and the many people on USENET who have given me names and leads.

Eamon de Valera, long-time Prime Minister and then President of the Republic of Ireland. A ringer: he was a mathematics professor before Irish independence.

MUSIC.

Ernst Ansermet, founder and conductor of the Orchestre de la Suisse Romande.

Pierre Boulez, Modernist composer and conductor.

Clifford Brown, Fifties jazz trumpeter.

Art Garfunkel, folk-rock singer. MA in mathematics from Columbia in 1967.

Worked on a PhD at Columbia, but chose to pursue his musical career instead.

Phillip Glass, composer, a Bachelor's from the University of Chicago.

Carole King, Sixties songwriter, and later a singer-songwriter. She dropped out after one year of college to pursue her music career.

Tom Lehrer, songwriter-parodist. PhD student in mathematics at Harvard.

Lawrence Leighton Smith, conductor and pianist.

THE OTHER ARTS.

Lewis Carroll, author of *Alice in Wonderland*, *Through the Looking Glass*, and other works. A ringer: he was a logician under his real name, Charles Lutwidge Dodgson.

Heloise (Poncé Cruse Evans), of *Hints from Heloise*. She minored in math.

Larry Niven, science fiction writer, winner of the Nebula and Hugo awards.

Omar Khayyam, author of *The Rubaiyat*. Another ringer: he published works on algebra and Euclid.

Alexander Solzhenitsyn, Nobel prize-winning novelist, a degree in mathematics and physics from the University of Rostov.

Bram Stoker, author of *Dracula*, took honors at Trinity University, Dublin.

Christopher Wren, the architect of St. Paul's Cathedral in London.

FINANCE.

John Maynard Keynes, the great economist. MA and 12th Wrangler, Cambridge University.

J. Pierpont Morgan, the banking, steel, and railroad magnate. Some of the Göttingen faculty tried to convince him to become a professional mathematician.

Ed Thorpe, one of the inventors of program-trading on Wall Street.

PHILOSOPHERS.

Edmund Husserl, the "Father of Phenomenology," PhD 1883 from Vienna.

Ludwig Wittgenstein, one of the giants of twentieth-century philosophy. Studied mathematical logic with Bertrand Russell.

ATHLETES AND OTHER COMPETITORS.

Michael Jordan, basketball superstar. He changed to another major in his junior year.

Davey Johnson, manager of the 1986 New York Mets.

Emanuel Lasker, world chess champion from 1894–1921. Another ringer, he was a mathematics professor with several published papers.

David Robinson, basketball star. BS in mathematics from Annapolis.

Frank Ryan, star quarterback for the Cleveland Browns in the sixties. PhD from Rice.

Virginia Wade, Wimbledon champion, BS in mathematics and physics from Sussex.

LITERARY CRIMINALS.

James Moriarty, former Professor of Mathematics, author of *The Dynamics of an Asteroid*, whose essay on the binomial theorem is said to have had a continental vogue, became the leader of the most sinister criminal conspiracy in Victorian England. He has been called “the Napoleon of Crime.” Sherlock Holmes’s nemesis.

*Mathematics Department
Lafayette College
Easton, PA 18042
buyskes@lafvax.lafayette.edu*

PICTURE PUZZLE (from the collection of Paul Halmos)



The smile came in 1984, soon after his great victory.
(see page 883.)

The Fundamental Theorem of Linear Algebra

Gilbert Strang

This paper is about a theorem and the pictures that go with it. The theorem describes the action of an m by n matrix. The matrix A produces a linear transformation from R^n to R^m —but this picture by itself is too large. The “truth” about $Ax = b$ is expressed in terms of four subspaces (two of R^n and two of R^m). The pictures aim to illustrate the action of A on those subspaces, in a way that students won’t forget.

The first step is to see Ax as a *combination of the columns of A* . Until then the multiplication Ax is just numbers. This step raises the viewpoint to subspaces. We see Ax in the *column space*. Solving $Ax = b$ means finding all combinations of the columns that produce b in the column space:

$$\left[\begin{array}{c|c|c|c|c} & & & \cdots & \end{array} \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1(\text{column 1}) + \cdots + x_n(\text{column } n) = b.$$

Columns of A

The column space is the range $R(A)$, a subspace of R^m . This abstraction, from entries in A or x or b to the picture based on subspaces, is absolutely essential. Note how subspaces enter *for a purpose*. We could invent vector spaces and construct bases at random. That misses the purpose. Virtually all algorithms and all applications of linear algebra are understood by moving to subspaces.

The key algorithm is elimination. Multiples of rows are subtracted from other rows (and rows are exchanged). There is no change in the *row space*. This subspace contains all combinations of the rows of A , which are the columns of A^T . The row space of A is the column space $R(A^T)$.

The other subspace of R^n is the *nullspace* $N(A)$. It contains all solutions to $Ax = 0$. Those solutions are not changed by elimination, whose purpose is to compute them. A by-product of elimination is to display the dimensions of these subspaces, which is the first part of the theorem.

The *Fundamental Theorem of Linear Algebra* has as many as four parts. Its presentation often stops with Part 1, but the reader is urged to include Part 2. (That is the only part we will prove—it is too valuable to miss. This is also as far as we go in teaching.) The last two parts, at the end of this paper, sharpen the first two. The complete picture shows the action of A on the four subspaces with the right bases. Those bases come from the singular value decomposition.

The Fundamental Theorem begins with

Part 1. *The dimensions of the subspaces.*

Part 2. *The orthogonality of the subspaces.*

The dimensions obey the most important laws of linear algebra:

$$\dim R(A) = \dim R(A^T) \quad \text{and} \quad \dim R(A) + \dim N(A) = n.$$

When the row space has dimension r , the nullspace has dimension $n - r$. Elimination identifies r pivot variables and $n - r$ free variables. Those variables correspond, in the echelon form, to columns with pivots and columns without pivots. They give the dimension count r and $n - r$. Students see this for the echelon matrix and believe it for A .

The *orthogonality* of those spaces is also essential, and very easy. Every x in the nullspace is perpendicular to every row of the matrix, exactly because $Ax = 0$:

$$Ax = \begin{bmatrix} -\text{row} & 1- \\ -\text{row} & 2- \\ -\text{row} & m- \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The first zero is the dot product of x with row 1. The last zero is the dot product with row m . One at a time, the rows are perpendicular to any x in the nullspace. So x is perpendicular to all combinations of the rows.

The nullspace $N(A)$ is orthogonal to the row space $R(A^T)$.

What is the fourth subspace? If the matrix A leads to $R(A)$ and $N(A)$, then its transpose must lead to $R(A^T)$ and $N(A^T)$. The fourth subspace is $N(A^T)$, ***the nullspace of A^T*** . We need it! The theory of linear algebra is bound up in the connections between row spaces and column spaces. If $R(A^T)$ is orthogonal to $N(A)$, then—*just by transposing*—the column space $R(A)$ is orthogonal to the “left nullspace” $N(A^T)$. Look at $A^T y = 0$:

$$A^T y = \begin{bmatrix} \text{column 1 of } A \\ \vdots \\ \text{column } n \text{ of } A \end{bmatrix} y = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since y is orthogonal to each column (producing each zero), y is orthogonal to the whole column space. The point is that A^T is just as good a matrix as A . Nothing is new, except A^T is n by m . Therefore the left nullspace has dimension $m - r$.

$A^T y = 0$ means the same as $y^T A = 0^T$. With the vector on the left, $y^T A$ is a combination of the *rows* of A . Contrast that with $Ax =$ combination of the columns.

The First Picture: Linear Equations

Figure 1 shows how A takes x into the column space. The nullspace goes to the zero vector. Nothing goes elsewhere in the left nullspace—which is waiting its turn.

With b in the column space, $Ax = b$ can be solved. There is a *particular* solution x_r in the row space. The *homogeneous* solutions x_n form the nullspace. The general solution is $x_r + x_n$. The particularity of x_r is that it is orthogonal to every x_n .

May I add a personal note about this figure? Many readers of *Linear Algebra and Its Applications* [4] have seen it as fundamental. It captures so much about $Ax = b$. Some letters suggested other ways to draw the orthogonal subspaces—artistically this is the hardest part. The four subspaces (and very possibly the figure itself) are of course not original. But as a key to the teaching of linear algebra, this illustration is a gold mine.

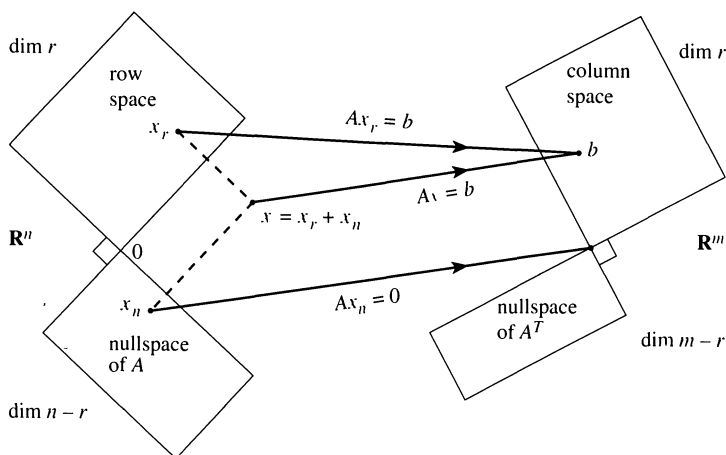


Figure 1. The action of A : Row space to column space, nullspace to zero.

Other writers made a further suggestion. They proposed a lower level textbook, recognizing that the range of students who need linear algebra (and the variety of preparation) is enormous. That new book contains Figures 1 and 2—also Figure 0, to show the dimensions first. The explanation is much more gradual than in this paper—but every course has to study subspaces! We should teach the important ones.

The Second Figure: Least Squares Equations

If b is not in the column space, $Ax = b$ cannot be solved. In practice we still have to come up with a “solution.” It is extremely common to have more equations than unknowns—more output data than input controls, more measurements than parameters to describe them. The data may lie close to a straight line $b = C + Dt$. A parabola $C + Dt + Et^2$ would come closer. Whether we use polynomials or sines and cosines or exponentials, the problem is still linear in the coefficients C, D, E :

$$\begin{array}{ccc} C + Dt_1 = b_1 & & C + Dt_1 + Et_1^2 = b_1 \\ \vdots & \text{or} & \vdots \\ C + Dt_m = b_m & & C + Dt_m + Et_m^2 = b_m \end{array}$$

There are $n = 2$ or $n = 3$ unknowns, and m is larger. There is no $x = (C, D)$ or $x = (C, D, E)$ that satisfies all m equations. $Ax = b$ has a solution only when the points lie exactly on a line or a parabola—then b is in the column space of the m by n matrix A .

The solution is to make the error $b - Ax$ as small as possible. Since Ax can never leave the column space, choose the closest point to b in that subspace. This point is the projection p . Then the error vector $e = b - p$ has minimal length.

To repeat: The best combination $p = A\bar{x}$ is the projection of b onto the column space. The error e is perpendicular to that subspace. Therefore $e = b - A\bar{x}$ is in the left nullspace:

$$A^T(b - A\bar{x}) = 0 \quad \text{or} \quad A^T A\bar{x} = A^T b.$$

Calculus reaches the same linear equations by minimizing the quadratic $\|b - Ax\|^2$. The chain rule just multiplies both sides of $Ax = b$ by A^T .

The “normal equations” are $A^T A \bar{x} = A^T b$. They illustrate what is almost invariably true—applications that start with a rectangular A end up computing with the square symmetric matrix $A^T A$. This matrix is invertible provided A has *independent columns*. We make that assumption: The nullspace of A contains only $x = 0$. (Then $A^T A x = 0$ implies $x^T A^T A x = 0$ which implies $Ax = 0$ which forces $x = 0$, so $A^T A$ is invertible.) The picture for least squares shows the action over on the right side—the splitting of b into $p + e$.

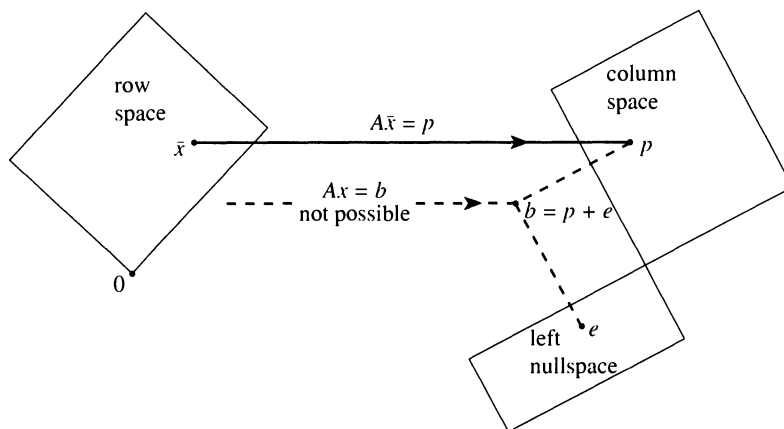


Figure 2. Least squares: \bar{x} minimizes $\|b - Ax\|^2$ by solving $A^T A \bar{x} = A^T b$.

The Third Figure: Orthogonal Bases

Up to this point, nothing was said about *bases for the four subspaces*. Those bases can be constructed from an echelon form—the output from elimination. This construction is simple, but the bases are not perfect. A really good choice, in fact a “canonical choice” that is close to unique, would achieve much more. To complete the Fundamental Theorem, we make two requirements:

Part 3. *The basis vectors are orthonormal.*

Part 4. *The matrix with respect to these bases is diagonal.*

If v_1, \dots, v_r is the basis for the row space and u_1, \dots, u_r is the basis for the column space, then $Av_i = \sigma_i u_i$. That gives a diagonal matrix Σ . We can further ensure that $\sigma_i > 0$.

Orthonormal bases are no problem—the Gram-Schmidt process is available. But a diagonal form involves eigenvalues. In this case they are the eigenvalues of $A^T A$ and AA^T . Those matrices are symmetric and positive semidefinite, so they have nonnegative eigenvalues and orthonormal eigenvectors (which are the bases!). Starting from $A^T A v_i = \sigma_i^2 v_i$, here are the key steps:

$$v_i^T A^T A v_i = \sigma_i^2 v_i^T v_i \quad \text{so that} \quad \|Av_i\| = \sigma_i$$

$$AA^T A v_i = \sigma_i^2 A v_i \quad \text{so that} \quad u_i = Av_i / \sigma_i \text{ is a unit eigenvector of } AA^T.$$

All these matrices have rank r . The r positive eigenvalues σ_i^2 give the diagonal entries σ_i of Σ .

The whole construction is called the *singular value decomposition (SVD)*. It amounts to a factorization of the original matrix A into $U\Sigma V^T$, where

1. U is an m by m orthogonal matrix. Its columns $u_1, \dots, u_r, \dots, u_m$ are basis vectors for the column space and left nullspace.
2. Σ is an m by n diagonal matrix. Its nonzero entries are $\sigma_1 > 0, \dots, \sigma_r > 0$.
3. V is an n by n orthogonal matrix. Its columns $v_1, \dots, v_r, \dots, v_n$ are basis vectors for the row space and nullspace.

The equations $Av_i = \sigma_i u_i$ mean that $AV = U\Sigma$. Then multiplication by V^T gives $A = U\Sigma V^T$.

When A itself is symmetric, its eigenvectors u_i make it diagonal: $A = U\Lambda U^T$. The singular value decomposition extends this spectral theorem to matrices that are not symmetric and not square. The eigenvalues are in Λ , the singular values are in Σ . The factorization $A = U\Sigma V^T$ joins $A = LU$ (elimination) and $A = QR$ (orthogonalization) as a beautifully direct statement of a central theorem in linear algebra.

The history of the *SVD* is cloudy, beginning with Beltrami and Jordan in the 1870's, but its importance is clear. For a very quick history and proof, and much more about its uses, please see [1]. "The most recurring theme in the book is the practical and theoretical value of this matrix decomposition." The *SVD* in linear algebra corresponds to the Cartan decomposition in Lie theory [3]. This is one more case, if further convincing is necessary, in which mathematics gets the properties right—and the applications follow.

Example

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = U\Sigma V^T.$$

All four subspaces are 1-dimensional. The columns of A are multiples of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ in U . The rows are multiples of $[1 \ 2]$ in V^T . Both $A^T A$ and AA^T have eigenvalues 50 and 0. So the only singular value is $\sigma_1 = \sqrt{50}$.

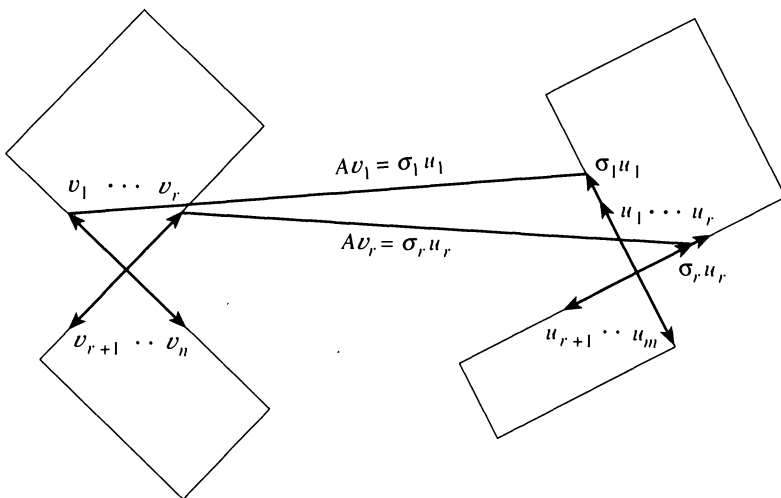


Figure 3. Orthonormal bases that diagonalize A .

The *SVD* expresses A as a combination of r rank-one matrices:

$$A = U\Sigma V^T = u_1\sigma_1v_1^T + \cdots + u_r\sigma_rv_r^T \quad \left(\text{here } A = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \right).$$

The Fourth Figure: The Pseudoinverse

The *SVD* leads directly to the “*pseudoinverse*” of A . This is needed, just as the least squares solution \bar{x} was needed, to invert A and solve $Ax = b$ when those steps are strictly speaking impossible. The pseudoinverse A^+ agrees with A^{-1} when A is invertible. The least squares solution of minimum length (having no nullspace component) is $x^+ = A^+b$. It coincides with \bar{x} when A has full column rank $r = n$ —then A^TA is invertible and Figure 4 becomes Figure 2.

A^+ takes the column space back to the row space [4]. On these spaces of equal dimension r , the matrix A is invertible and A^+ inverts it. On the left nullspace, A^+ is zero. I hope you will feel, after looking at Figure 4, that this is the one natural best definition of an inverse. Despite those good adjectives, the *SVD* and A^+ is too much for an introductory linear algebra course. It belongs in a second course. Still the picture with the four subspaces is absolutely intuitive.

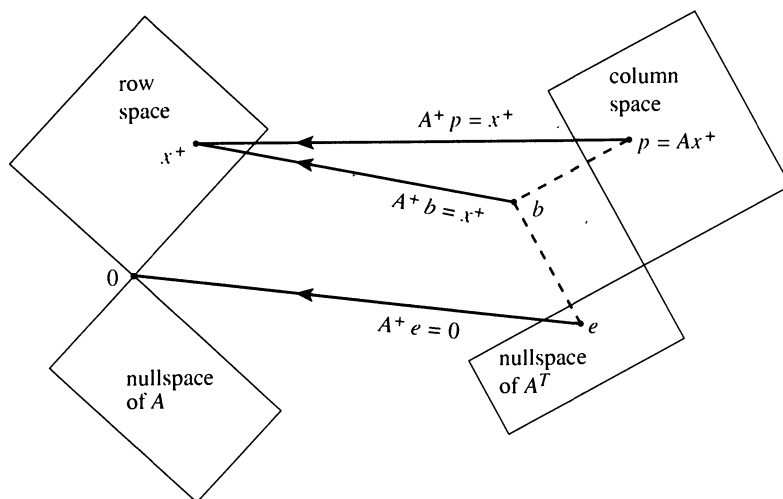


Figure 4. The inverse of A (where possible) is the pseudoinverse A^+ .

The *SVD* gives an easy formula for A^+ , because it chooses the right bases. Since $Av_i = \sigma_i u_i$, the inverse has to be $A^+u_i = v_i/\sigma_i$. Thus the pseudoinverse of Σ contains the reciprocals $1/\sigma_i$. The orthogonal matrices U and V^T are inverted by U^T and V . All together, the pseudoinverse of $A = U\Sigma V^T$ is $A^+ = V\Sigma^+U^T$.

Example (continued)

$$A^+ = \frac{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}{\sqrt{5}} \begin{bmatrix} 1/\sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}}{\sqrt{10}} = \frac{1}{50} \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$$

Always A^+A is the identity matrix on the row space, and zero on the nullspace:

$$A^+A = \frac{1}{50} \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix} = \text{projection onto the line through } \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Similarly AA^+ is the identity on the column space, and zero on the left nullspace:

$$AA^+ = \frac{1}{50} \begin{bmatrix} 5 & 15 \\ 15 & 45 \end{bmatrix} = \text{projection onto the line through } \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

A Summary of the Key Ideas

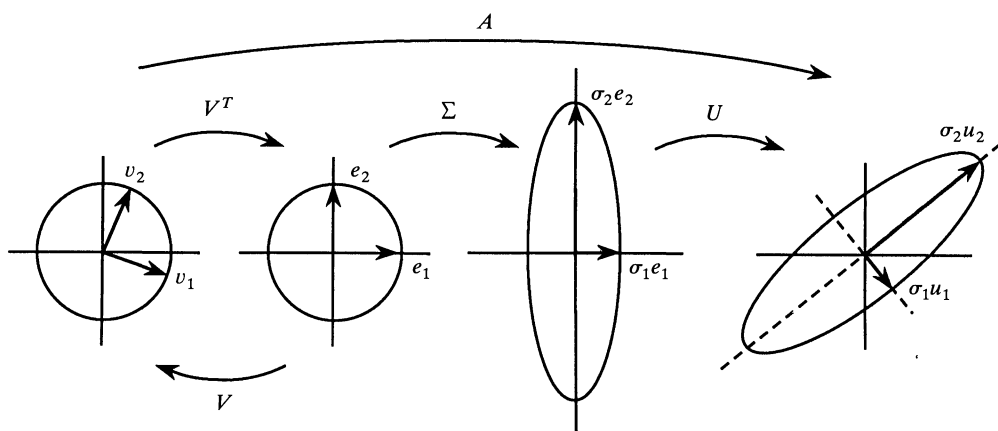
From its r -dimensional row space to its r -dimensional column space, A yields an invertible linear transformation.

Proof: Suppose x and x' are in the row space, and Ax equals Ax' in the column space. Then $x - x'$ is in both the row space and nullspace. It is perpendicular to itself. Therefore $x = x'$ and the transformation is one-to-one.

The SVD chooses good bases for those subspaces. Compare with the Jordan form for a real square matrix. There we are choosing the *same basis* for both domain and range—our hands are tied. The best we can do is $SAS^{-1} = J$ or $SA = JS$. In general J is not real. If real, then in general it is not diagonal. If diagonal, then in general S is not orthogonal. By choosing *two bases*, not one, every matrix does as well as a symmetric matrix. The bases are orthonormal and A is diagonalized.

Some applications permit two bases and others don't. For powers A^n we need S^{-1} to cancel S . Only a similarity is allowed (one basis). In a differential equation $u' = Au$, we can make one change of variable $u = Sv$. Then $v' = S^{-1}ASv$. But for $Ax = b$, the domain and range are philosophically "not the same space." The row and column spaces are isomorphic, but their bases can be different. And for least squares the SVD is perfect.

This figure by Tom Hern and Cliff Long [2] shows the diagonalization of A . Basis vectors go to basis vectors (principal axes). A circle goes to an ellipse. The matrix is factored into $U\Sigma V^T$. Behind the scenes are *two* symmetric matrices A^TA and AA^T . So we reach two orthogonal matrices U and V .



We close by summarizing the action of A and A^T and A^+ :

$$Av_i = \sigma_i u_i \quad A^T u_i = \sigma_i v_i \quad A^+ u_i = v_i / \sigma_i \quad 1 \leq i \leq r.$$

The nullspaces go to zero. Linearity does the rest.

The support of the National Science Foundation (DMS 90-06220) is gratefully acknowledged.

REFERENCES

1. Gene Golub and Charles Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins University Press (1989).
2. Thomas Hern and Cliff Long, Viewing some concepts and applications in linear algebra, *Visualization in Teaching and Learning Mathematics*, MAA Notes 19 (1991) 173–190.
3. Roger Howe, Very basic Lie theory, *American Mathematical Monthly*, 90 (1983) 600–623.
4. Gilbert Strang, *Linear Algebra and Its Applications*, 3rd ed., Harcourt Brace Jovanovich (1988).
5. Gilbert Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press (1993).

Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139
gs@math.mit.edu

An Identity of Daubechies

The generalization of an identity of Daubechies using a probabilistic interpretation by D. Zeilberger [100 (1993) 487], has already appeared in SIAM Review Problem 85-10 (June, 1985) in a slightly more general context. In addition to a similar probabilistic derivation there is also a direct algebraic proof. Incidentally, problem 10223 [99 (1992) 462] is the same as the identity of Daubechies and a slight generalization of this identity has appeared previously as problem 183, *Crux Math.* 3(1977) 69–70 and came from a list of problems considered for the Canadian Mathematical Olympiad. There was an inductive solution of the latter by Mark Kleinman, a high school student at the time and one of the top students in the U.S.A.M.O. and the I.M.O.

M. S. Klamkin
Department of Mathematics
University of Alberta
Edmonton, Alberta
CANADA T6G 2G1

A Simple Proof of the Jordan-Alexander Complement Theorem

Albrecht Dold

The complements of homeomorphic subsets $A, B \subset \mathbb{R}^n$ of Euclidean space need not be homeomorphic, $A \approx B \not\Rightarrow (\mathbb{R}^n - A) \approx (\mathbb{R}^n - B)$. This is well illustrated by classical knot theory, i.e. when A, B are knots in \mathbb{R}^3 . The complements usually have different fundamental groups in this case, $\pi_1(\mathbb{R}^3 - A) \not\cong \pi_1(\mathbb{R}^3 - B)$, and this fundamental group serves to distinguish non-equivalent knots.

On the other hand, it is a classical consequence of **Alexander** duality (cf. [D], VIII, 8.15) that the homology groups of the complements agree if A, B are homeomorphic closed subsets of \mathbb{R}^n . Thus,

Theorem. *If $A, B \subset \mathbb{R}^n$ are homeomorphic closed subsets then their complements have isomorphic homology groups, $H(\mathbb{R}^n - A) \cong H(\mathbb{R}^n - B)$,—also in generalised (co-)homology.*

If the coefficients of homology are taken in a commutative ring R with 1 then the rank of $H_0(\mathbb{R}^n - A)$ equals the number of components of $\mathbb{R}^n - A$ (almost by definition of H_0). Therefore,

Corollary. *The complements of homeomorphic closed subsets $A, B \subset \mathbb{R}^n$ have the same number of components.*

If $A = \{x \in \mathbb{R}^n \mid \|x\| = 1\} = S^{n-1}$, the **Jordan** separation theorem is: Every subset $B \subset \mathbb{R}^n$, $n > 1$, which is homeomorphic to S^{n-1} separates \mathbb{R}^n into two regions.

In this note we give a simple proof of the theorem. It uses basic properties only of homology, namely homotopy invariance and Mayer-Vietoris sequences of open subsets of Euclidean spaces. The reader might take singular or simplicial homology but the proof also works in general (co-)homology—no dimension axiom is required. No priority is claimed for this note. Its methods are familiar in topology and algebraic geometry; the intention is to publicize an elegant argument.

It is convenient to use reduced homology in the proof. The *reduced homology* $\tilde{H}X$ of a non-empty space X is the kernel of the homomorphism $HX \rightarrow HP$ which is induced by the map $X \rightarrow P$ onto the one-point space P . If we choose a point in X , which we write as a map $P \rightarrow X$, then the composition $P \rightarrow X \rightarrow P$ is the identity map, hence $HX = \text{im}(HP \rightarrow HX) \oplus \ker(HX \rightarrow HP) = HP \oplus \tilde{H}X$. Thus, HX differs from $\tilde{H}X$ by the constant summand HP only. In particular, $\tilde{H}P = 0$ and by homotopy invariance, $\tilde{H}X = 0$ for every contractible space X .

Proposition 1. *For every closed subset $A \subset \mathbb{R}^n$, $A \neq \mathbb{R}^n$, we have $\tilde{H}_i(\mathbb{R}^n - A) \cong \tilde{H}_{i+1}(\mathbb{R}^{n+1} - A)$, where $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$, $\mathbb{R}^n = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+1}$.*

Proof: Put $Z = \mathbb{R}^{n+1} - A$,

$$Z_+ = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} | t > 0, \text{ or } x \in (\mathbb{R}^n - A)\},$$

$$Z_- = \{(x, t) \in \mathbb{R}^{n+1} | t < 0 \text{ or } x \in (\mathbb{R}^n - A)\}.$$

Then Z_+, Z_- are open in Z ,

$$Z_+ \cup Z_- = Z, Z_+ \cap Z_- = (\mathbb{R}^n - A) \times \mathbb{R}.$$

Furthermore, Z_+ and Z_- are contractible (the deformation $(x, t) \mapsto (x, (1 - \tau)t + \tau)$, $0 \leq \tau \leq 1$, moves Z_+ into the hyperplane $t = 1$ which in turn deforms into a point), hence $\tilde{H}(Z_+) = 0 = \tilde{H}(Z_-)$. The reduced Mayer-Vietoris sequence (which is the ordinary Mayer-Vietoris sequence without the superfluous constant summands HP ; cf. [D], III, 8.15) has the form

$$\begin{aligned} \tilde{H}_{i+1}(Z_+) \oplus \tilde{H}_{i+1}(Z_-) &\rightarrow \tilde{H}_{i+1}(Z_+ \cup Z_-) \rightarrow \tilde{H}_i(Z_+ \cap Z_-) \\ &\rightarrow \tilde{H}_i(Z_+) \oplus \tilde{H}_i(Z_-). \end{aligned}$$

As it is exact and $\tilde{H}(Z_+) = 0 = \tilde{H}(Z_-)$ it amounts to an isomorphism $\tilde{H}_{i+1}(Z_+ \cup Z_-) \cong \tilde{H}_i(Z_+ \cap Z_-)$. But $Z_+ \cup Z_- = \mathbb{R}^{n+1} - A$, and $Z_+ \cap Z_- = (\mathbb{R}^n - A) \times \mathbb{R}$; the latter deforms into $(\mathbb{R}^n - A) \times \{0\} = \mathbb{R}^n - A$. ■

Iterating Proposition 1 we get

Proposition 2. *For every closed subset $A \subset \mathbb{R}^n$, $A \neq \mathbb{R}^n$, and every $q \geq 0$, $\tilde{H}_{i+q}(\mathbb{R}^{n+q} - A) \cong \tilde{H}_i(\mathbb{R}^n - A)$.* ■

Proposition 3. *If $A \subset \mathbb{R}^p$, $B \subset \mathbb{R}^q$ are closed subsets which are homeomorphic then the complements of $A = A \times \{0\}$ and $B = \{0\} \times B$ in $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$ are also homeomorphic, $(\mathbb{R}^{p+q} - A) \approx (\mathbb{R}^{p+q} - B)$.*

Proof: (Compare [FM], §3). Let $\varphi: A \rightleftharpoons B: \psi$ be reciprocal homeomorphisms, $\psi\varphi = 1_A$, $\varphi\psi = 1_B$. By Tietze's Lemma, these extend to continuous maps $\Phi: \mathbb{R}^p \rightleftharpoons \mathbb{R}^q: \Psi$. The maps $L, R: \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}^p \times \mathbb{R}^q$, $L((x, y) = (x, y - \Phi(x))$, $R(x, y) = (x - \Psi(y), y)$ are self-homeomorphisms of \mathbb{R}^{p+q} , and they map the graph $\Gamma = \{(x, y) \in \mathbb{R}^p \times \mathbb{R}^q | x \in A, y = \varphi(x)\} = \{(x, y) | y \in B, x = \psi(y)\}$ onto A resp. B . Hence $(\mathbb{R}^{p+q} - A) \stackrel{L}{\cong} (\mathbb{R}^{p+q} - \Gamma) \stackrel{R}{\cong} (\mathbb{R}^{p+q} - B)$. ■

Proof of the theorem. If both A and B are $\neq \mathbb{R}^n$ we apply propositions 2, 3, 2 in this order, $\tilde{H}_i(\mathbb{R}^n - A) \cong \tilde{H}_{i+n}(\mathbb{R}^{n+n} - A) \cong \tilde{H}_{i+n}(\mathbb{R}^{n+n} - B) \cong \tilde{H}_i(\mathbb{R}^n - B)$. Adding $H_i P$ to both ends gives $H_i(\mathbb{R}^n - A) \cong H_i(\mathbb{R}^n - B)$, as required.

If $A = \mathbb{R}^n$ we still have $\tilde{H}_j(\mathbb{R}^{n+1} - A) \cong \tilde{H}_j(\mathbb{R}^{n+1} - B)$ by the same argument; in particular, $\tilde{H}_0(\mathbb{R}^{n+1} - A) \cong \tilde{H}_0(\mathbb{R}^{n+1} - B)$. But $\mathbb{R}^{n+1} - A = \mathbb{R}^{n+1} - \mathbb{R}^n$ has two components and $\mathbb{R}^{n+1} - B$ has only one component—unless $B = \mathbb{R}^n$. Therefore, (in ordinary homology), $\text{rank}(\tilde{H}_0(\mathbb{R}^{n+1} - A)) = 1$ and $\text{rank}(\tilde{H}_0(\mathbb{R}^{n+1} - B)) = 0$ —unless $B = \mathbb{R}^n$. Therefore, $B = \mathbb{R}^n = A$. ■

REFERENCES

- [D] Dold, A., Lectures on Algebraic Topology. Springer Verlag Berlin-Heidelberg-New York, 1972/80.
 [FM] Fulton, W., MacPherson, R., Categorical Framework for the Study of Singular Spaces. Mem. Amer. Math. Soc. 243 (1981).

*Mathematisches Institut
 Im Neuenheimer Feld 288
 D-69120 Heidelberg 1
 Germany*

Squaring the Circle with Holes

Hansklaus Rummler

1. WALLIS' PRODUCT. Among the approximations of π , Wallis' product

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots$$

is perhaps the most fascinating one. Sure, it is not really useful in calculating π , the product converging very slowly. But the formula is already interesting for its history: Wallis' somewhat mysterious—or even mystic—discovery of the formula inspired Newton to similar calculations, leading finally to the binomial series (see [1]).

Nowadays, the proof of Wallis' formula has become a standard exercise: Calculating the integral

$$I_m = \int_0^{\pi/2} \sin^m x \, dx$$

for every natural number m leads to

$$I_{2n} = \frac{\pi}{2} \cdot \frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \quad \text{and} \quad I_{2n+1} = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \frac{2n}{2n+1},$$

and from this Wallis' product formula is easily derived.

An alternative proof is obtained by taking $z = \frac{1}{2}$ in the Weierstraß product

$$\sin(\pi z) = \pi z \prod_{\nu=1}^{\infty} \left(1 - \frac{z^2}{\nu^2}\right).$$

Unfortunately, neither proof helps to understand the formula. To explain what we mean by *understanding a formula*, let us consider Vieta's formula:

$$\begin{aligned} \frac{2}{\pi} = & \sqrt{\frac{1}{2}} + \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}} \cdot \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}}} \\ & \cdot \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}}}} \cdots \end{aligned}$$

The factors of this infinite product are much more complicated than those of Wallis' product, but they have a simple geometric meaning, because they represent length ratios: If l_n denotes the length of a regular 2^n -gon inscribed in the unit circle, it can be shown that the factors of Vieta's product are just the ratios

$l_n : l_{n+1}$, and the formula is immediately clear:

$$\frac{2}{\pi} = \frac{l_1}{l_\infty} = \lim_{n \rightarrow \infty} \frac{l_1}{l_2} \cdot \frac{l_2}{l_3} \cdots \frac{l_n}{l_{n+1}} = \frac{l_1}{l_2} \cdot \frac{l_2}{l_3} \cdot \frac{l_3}{l_4} \cdot \frac{l_4}{l_5} \cdots$$

(The inscribed 2-gon is a diameter, counted twice.)

2. WALLIS' SIEVE. Instead of trying to *understand* Wallis' product formula in the same sense we understand Vieta's formula, we shall *interpret* it, constructing a subset of the unit square that is easily seen to have area $\frac{8}{9} \cdot \frac{24}{25} \cdot \frac{48}{49} \cdots$, which, by Wallis' product formula, is just the area of the inscribed disk, namely $\pi/4$.

In order to construct this set, let us say that we *punch a hole of order n* into a square, n being an odd integer, if we take away the middle open one of the n^2 congruent small squares into which we can decompose the given square.

Now take a compact unit square and punch a hole of order 3 into this square. The remaining set W_1 has of course area

$$\mu(W_1) = \frac{8}{9}.$$

Punching a hole of order 5 into each of the 8 small squares forming W_1 , we get a set W_2 consisting of $8 \cdot 24$ small squares and with area

$$\mu(W_2) = \frac{8}{9} \cdot \frac{24}{25}.$$

Continuing in this way by punching holes of order 7, 9, 11 and so on, we get finally *Wallis's sieve*, a compact set W_∞ with area

$$\mu(W_\infty) = \frac{8}{9} \cdot \frac{24}{25} \cdot \frac{48}{49} \cdots = \frac{\pi}{4}.$$

The following figures show the first three steps of our construction:

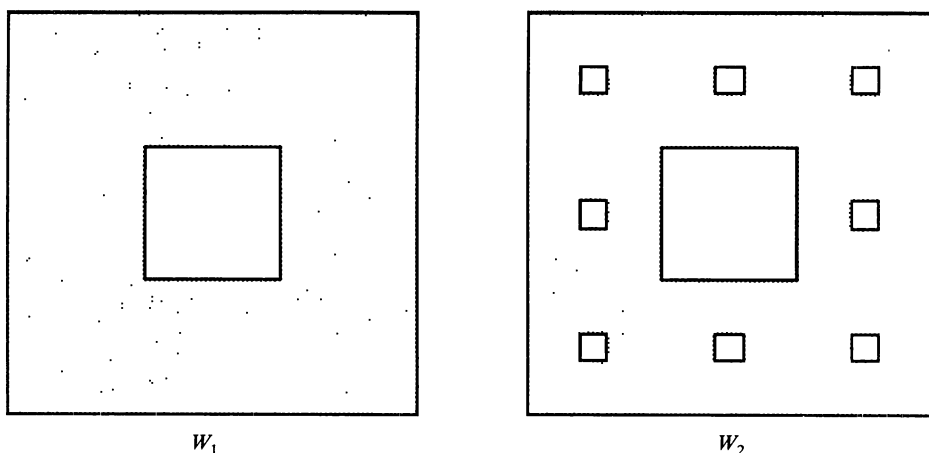
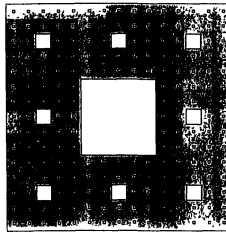


Figure 1



W_3

Figure 2

3. WALLIS' SIEVE AND LEBESGUE MEASURE. So far, there seems to be no problem in calculating the area of Wallis' sieve W_∞ , and by construction this area is just $\mu(W_\infty) = \frac{8}{9} \cdot \frac{24}{25} \cdot \frac{48}{49} \cdots = \pi/4$. But we have to be careful: *area* here means *Lebesgue measure*, because W_∞ is not measurable in Jordan's sense, its interior measure being 0. W_∞ does not even contain any product set $A \times B$ with $A, B \subset \mathbb{R}$ having positive Lebesgue measure. To see this, consider a maximal product subset of W_1 , for instance $[0, 1] \times ([0, \frac{1}{3}] \cup [\frac{2}{3}, 1])$. This subset has measure $\frac{2}{3}$, a maximal product subset of W_2 has measure $\frac{2}{3} \cdot \frac{4}{5}$, and so on. Therefore, a maximal product subset of W_∞ has measure $\frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots = 0$.

Thus, Wallis' sieve W_∞ is an example of a subset of the plane \mathbb{R}^2 with positive Lebesgue measure, but not admitting any product subset with positive Lebesgue measure.

REFERENCE

1. C. H. Edwards, Jr., *The Historical Development of the Calculus*, Springer-Verlag, New York, 1979, pp. 166–176.

*Institut de Mathématiques
de l'Université de Fribourg,
Chemin du Musée 23,
CH-1700 FRIBOURG*

Fermat's Last Problem

An Englishman named Wiles discovered the key,
To Fermat's Last Problem using geometry.
By proving the sum of two powers,
Is a number to the power.
If and only if the power is smaller than three.

—Nats Wolraf

NOTES

Edited by: John Duncan

Simplifying the Proof of Dirichlet's Theorem

Paul Monsky

Dirichlet showed that an arithmetic progression $a, a + D, a + 2D, \dots$ with $D \geq 1$ and $(a, D) = 1$ contains infinitely many primes. Most of his argument is accessible to undergraduate mathematics majors, but a proof of the theorem is seldom presented to them because of the reputed difficulty of a key step—showing that certain infinite sums are non-zero. This note outlines a simple proof of the non-vanishing of these sums. The argument is very close to one given by Gelfond, [1], but is easier and works well in the classroom.

The sums I'll treat may be described as follows. A “character to the modulus D ” is a function $\chi: \mathbb{Z} \rightarrow \mathbb{C}$ satisfying:

- (1) If $a \equiv b(D)$, then $\chi(a) = \chi(b)$
- (2) $\chi(ab) = \chi(a)\chi(b)$
- (3) $\chi(a) = 0$ if and only if $(a, D) > 1$.

χ is said to be *real* if it takes real values (which can only be 1, -1 , or 0), *non-principal* if it takes values other than 0 or 1. Suppose for example that D is an odd prime. Then the “Legendre symbol”, taking each quadratic residue of D to 1, each non-residue to -1 and each multiple of D to 0 is a real non-principal character. For a non-principal χ , $\sum_1^\infty \chi(n)/n$ converges. (This follows from summation by parts; see the argument given in the last paragraph of this note.)

The usual approach to proving Dirichlet's theorem involves several standard analytic techniques (see [2], for example); the main non-formal step is showing that $\sum_1^\infty \chi(n)/n \neq 0$ whenever χ is real and non-principal (the result is also needed for non-real χ , but this is fairly easily handled). Dirichlet's original non-vanishing proof involved a detour through the theory of binary quadratic forms. Modern proofs generally use ideal theory in quadratic number fields or some complex variable theory. Elementary proofs are also known, but are more complicated than the one I'll now present.

One begins by defining c_n to be $\sum \chi(d)$ where d ranges over the positive divisors of n . Evidently $c_{p^a} = 1 + \chi(p) + \chi(p)^2 + \dots + \chi(p)^a \geq 0$. It follows easily that $c_n \geq 0$ for all n . Furthermore, $c_n = 1$ whenever n is a power of a prime p dividing D . In particular $\sum_1^\infty c_n = \infty$.

Next, following [1], one sets $f(t) = \sum_1^\infty \chi(n)t^n/(1 - t^n)$. The series evidently converges in $[0, 1)$. Expanding each $t^n/(1 - t^n)$ one finds that $f(t) = \sum_1^\infty c_n t^n$. The paragraph above shows that $f(t) \rightarrow \infty$ as $t \rightarrow 1^-$. Suppose now that $\sum_1^\infty \chi(n)/n = 0$. Then $-f(t) = \sum_1^\infty \chi(n)[1/n(1 - t)] - \{t^n/(1 - t^n)\}$; write this as $\sum_1^\infty \chi(n)b_n$.

The critical observation is that $b_1 \geq b_2 \geq b_3 \geq \dots$. Note first that

$$\begin{aligned} (1 - t)(b_n - b_{n+1}) &= \frac{1}{n} - \frac{1}{n+1} - \frac{t^n}{1 + t + \dots + t^{n-1}} + \frac{t^{n+1}}{1 + t + \dots + t^n} \\ &= \frac{1}{n(n+1)} - \frac{t^n}{(1 + t + \dots + t^{n-1})(1 + t + \dots + t^n)}. \end{aligned}$$

Since $(1 + t + \cdots + t^{n-1}) \geq nt^{(n-1)/2} \geq nt^{n/2}$ while $(1 + t + \cdots + t^n) \geq (n + 1)t^{n/2}$ (this is the inequality of the arithmetic and geometric mean), $b_n - b_{n+1} \geq 0$.

Now χ is periodic of period D , and $\sum_1^D \chi(n) = 0$. So the numbers $\chi(1), \chi(1) + \chi(2), \chi(1) + \chi(2) + \chi(3), \dots$ are bounded in absolute value by D . Since $b_n \searrow 0$, the standard Abel rearrangement of the infinite sum $\sum_1^\infty \chi(n)b_n$ shows that $|\sum_1^\infty \chi(n)b_n| \leq Db_1 = D$, contradicting the unboundedness of f on $[0, 1]$.

REFERENCES

1. A. O. Gelfond, *Elementary Methods in Analytic Number Theory*, Rand McNally and Co, 1965, pp. 47–49.
2. H. Davenport, *Multiplicative Number Theory*, Markham, Chicago, 1967.

Department of Mathematics
Brandeis University
Waltham, MA 02254-9110

Why is P^2 Not Embeddable in R^3 ?

Hiroshi Maehara

The projective plane P^2 is the closed surface obtained by pasting a Möbius band and a 2-cell together along their boundaries. The surface P^2 is not embeddable in the 3-dimensional Euclidean space R^3 . Though this fact is well known, no handy proof seems to be furnished yet. (A proof in Spanier [3], for instance, requires cohomology theory.) Here, we offer a short and clear-cut proof of the non-embeddability of P^2 in R^3 by applying the *Link Appearing Theorem*.

Our figures in R^3 are assumed to be *tame*. (A figure X in R^3 is tame if there exists a homeomorphism $f: R^3 \rightarrow R^3$ such that $f(X)$ is a polygonal or polyhedral figure.) Thus, we consider only tame embeddings. A (2-component) *link* is an embedding of a pair of circles in R^3 . Let us call a link *trivial* if one of the two

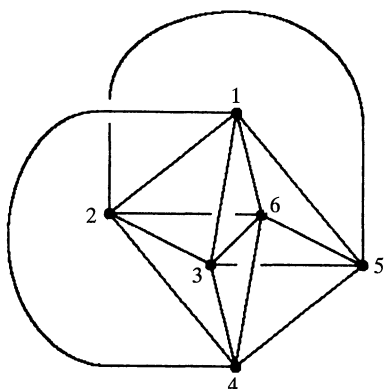


Figure 1.

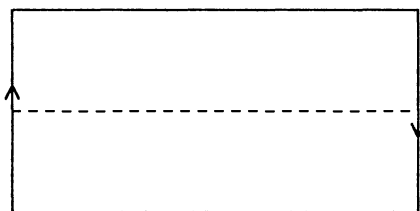


Figure 2.

curves bounds a 2-cell in R^3 that is disjoint from the other curve. Otherwise, it is *non-trivial*.

Now, consider a set of six points in R^3 , and assume that each pair of these points is connected by a simple curve such that the curves meet only at their endpoints. Such a figure is called a *complete 6-graph* and is usually denoted K_6 . Fig. 1 shows a K_6 in which six points are indicated by 1, 2, ..., 6. A simple closed curve in a K_6 is called a *cycle* of the K_6 . We indicate a cycle by a sequence of points in the order of appearing when we trace the cycle. In the K_6 of Fig. 1, the pair of cycles 135 and 246 forms a non-trivial link.

Link Appearing Theorem. *Any complete 6-graph in R^3 contains a pair of disjoint cycles that forms a non-trivial link.* ■

This theorem was proved by Sachs [2] and independently by Conway-Gordon [1]. Its proof is not difficult, see [1] or [2] for the detail.

In a rectangular representation of a Möbius band M , the line segment connecting the midpoints of the to-be-identified sides (the dotted line in Fig. 2) represents a simple closed curve in the Möbius band M . This closed curve is called the *meridian* of M .

Lemma. *For any embedding of a Möbius band M in R^3 , the pair $(\partial M, C)$ of the boundary ∂M and the meridian C of M forms a non-trivial link.*

Proof: Consider the K_6 on the Möbius band M represented in Fig. 3. (Each pair of the six points 1, 2, ..., 6 is, indeed, connected by a simple curve. For example, the line segment from the point 1 to the right-top e and the line segment from the left-bottom e to the point 3 make together a simple curve connecting 1 and 3.) This K_6 contains ten pairs of disjoint cycles:

$$(\underline{123}, \underline{456}), (\underline{124}, \underline{356}), (\underline{125}, \underline{346}), (\underline{126}, \underline{345}), (\underline{134}, \underline{256}),$$

$$(\underline{135}, \underline{246}), (\underline{136}, \underline{245}), (\underline{145}, \underline{236}), (\underline{146}, \underline{235}), (\underline{156}, \underline{234}).$$

Each underlined cycle bounds a 2-cell in M that is disjoint from its partner cycle. For example, the cycle 135 bounds the 2-cell shaded in Fig. 3. Hence, in any embedding of the Möbius band M in R^3 , nine pairs of cycles of K_6 other than $(134, 256)$ are trivial links. Therefore, $(134, 256)$ must be a non-trivial link by the Link Appearing Theorem. The cycle 256 is the meridian of M , and the cycle 134 is the boundary ∂M . ■

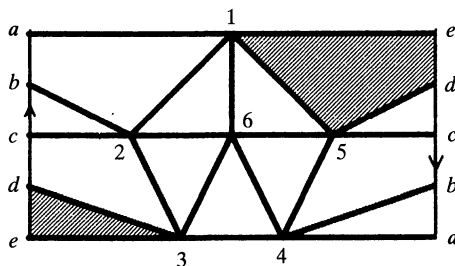


Figure 3.

Proof of the non-embeddability of P^2 in R^3 . Suppose P^2 is embedded in R^3 . By removing an open 2-cell D from the surface P^2 , we have a Möbius band M . Then, the boundary ∂M and the meridian C of M form together a non-trivial link. Therefore, C and the 2-cell D must intersect each other, a contradiction. ■

REFERENCES

1. J. H. Conway and C. McA. Gordon, Knots and links in spatial graphs, *J. Graph Theory*, 7 (1985), 445–452.
2. H. Sachs, On a spatial analogue of Kuratowski's theorem on planar graphs—an open problem, in *Graph Theory*, Lagow 1981, Lecture Note in Math. #1018, Springer Verlag, Berlin, 1982, 230–241.
3. E. H. Spanier, *Algebraic topology*, McGraw-Hill Book Company, New York, 1966.

*College of Education
Ryukyu University
Nishihara, Okinawa, Japan*

Polynomial Root Dragging

Bruce Anderson

1. INTRODUCTION. Rolle's Theorem and other results (such as those found in M. Marden [1] and anthologized in E. Barbeau [2]) furnish insight about the location of the zeros of the derivative of a polynomial (i.e. the critical points) relative to the location of the zeros of the polynomial. These results tend to be “static” in that they indicate where the critical points should be expected within certain bounds defined by the fixed location of the roots of $p(x)$ (e.g. within intervals bounded by the roots of the polynomial for real roots, or within a complex hull for the complex case). This paper will, in contrast, explore a simple “dynamic” result, showing how the roots of the derivative will be “affected” as we move (or drag) the roots of the polynomial, provided all the roots are real. The results are given in Theorem 2.1 and Corollary 2.2. We then show that this result does not generalize to complex roots in the obvious way.

The root dragging result of Corollary 2.2 is then employed to address the questions: Do the quartic polynomials produce all possible arrangements of critical points which satisfy Rolle's theorem? Or are there additional constraints on the possible arrangement of real critical points for quartics? Theorem 3.1 will furnish the perhaps surprising answer.

2. ROOT DRAGGING. Let $p(x)$ be a polynomial of degree n with all real distinct roots $x_1 < x_2 < \cdots < x_n$. Suppose we “drag to the right” some or all of these roots. I.e. we construct a new n th degree polynomial q with all real distinct roots $x'_1 < x'_2 < \cdots < x'_n$ such that $x'_i > x_i$ for all integers i between 1 and n . The derivatives of p and q , which of course are polynomials of degree $n - 1$, must also have all real distinct roots from Rolle's theorem. Let $z_1 < z_2 < \cdots < z_{n-1}$ and $z'_1 < z'_2 < \cdots < z'_{n-1}$ be the roots of p' and q' , respectively. (By Rolle's theorem, $x_k < z_k < x_{k+1}$ and $x'_k < z_k < x'_{k+1}$ for all integers k between 1 and $n - 1$).

Theorem 2.1. (Root Dragging Theorem). *The roots of q' will each be to the right of the corresponding roots of p' ; i.e. $z'_k > z_k$ for all integers k between 1 and $n - 1$.*

Proof: Our analysis will be in the spirit of the proof of the Gauss-Lucas Theorem found in Marden [1]. We suppose there is some k such that the corresponding roots, z_k and z'_k , are *not* in the order guaranteed by the theorem, i.e. $z'_k < z_k$. We show this leads to a contradiction. As shown in Marden [1], we know that the root z_k of p' must satisfy the equation:

$$\sum_{i=1}^n \frac{1}{z_k - x_i} = 0. \quad (1)$$

Likewise, the root z'_k of q' must satisfy:

$$\sum_{i=1}^n \frac{1}{z'_k - x'_i} = 0. \quad (2)$$

But since $x'_i > x_i$ and $z'_k < z_k$ (by assumption) we conclude:

$$z'_k - x'_i < z_k - x_i \quad (3)$$

Now, since z_k lies between x_k and x_{k+1} and z'_k lies between x'_k and x'_{k+1} , both sides of inequality (3) will be of the same sign. Thus

$$\frac{1}{z'_k - x'_i} > \frac{1}{z_k - x_i}. \quad (4)$$

Since this is true for all i , sums (1) and (2) cannot both equal zero. Q.E.D.

Corollary 2.2. *Let p and q be the same polynomials described in the theorem above. The roots of any derivative $q^{(j)}$ will each be to the right of the corresponding roots of $p^{(j)}$. I.e. if we shift roots of p to the right, the roots of all its derivatives will also shift to the right.*

Proof: Follows easily by induction on j . Q.E.D.

Remarks 2.3. (i) With a little care the requirement that the roots be distinct (i.e. no multiple roots) may be dropped. (ii) Essentially Corollary 2.2 says that the roots of the derivatives of a polynomial “follow” the roots of the polynomial (assuming all the roots are real). A more refined analysis which will not be presented here gives the following result: The roots of the derivatives will all move faster than the slowest moving root of the polynomial and slower than the fastest moving root of the polynomial.

3. APPLICATION OF THE ROOT DRAGGING THEOREM. Let $p(x)$ be a fourth degree polynomial whose (four) roots are all real and distinct. Call the inner two roots a_1 and a_2 . Now by Rolle's theorem, $p'(x)$ must have exactly three real distinct roots. Call the middle root b . Iterating Rolle's, $p''(x)$ must have two real distinct roots (which we call c_1 and c_2), and $p^{(3)}$ must have one real root, d . By elementary analysis, d will be the average value of the four roots of $p(x)$.

Theorem 3.1 (Unconstructible fourth degree polynomial). *If $a_1 < c_1$ and $a_2 < c_2$ then $b < d$.*

Remarks 3.2. Figure 1 illustrates the arrangement of roots which Theorem 3.1 states is unconstructible. Here “0” represents the location along the real number line of a root of p , “1” represents a root of p' , and so on).

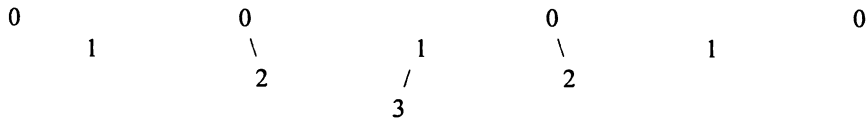


Figure 1. Theorem 3.1 states that this arrangement of roots is unconstructible, where “0” represents the location of a root of $p(x)$, “1” represents the location of a root of $p'(x)$ and so on.

Proof of Theorem 3.1: Begin shifting the right-most “0” to the right. Since the location of the “3” is the average of the “0’s” and since the middle “1” must lie between the second and third “0’s” by Rolle’s Theorem, the “3” must eventually line up with the middle “1” as we continue shifting the right-most “0” to the right. Meanwhile, by Corollary 2.2, the 2’s must shift to the right. Thus if the polynomial represented by Figure 1 is constructible, then so must Figure 2.

This means the polynomial must be symmetric around the middle “1”, since we have a fourth degree polynomial with the first and third derivatives equal to zero there. But clearly the ordering depicted in Figure 2 is not symmetric. Q.E.D.

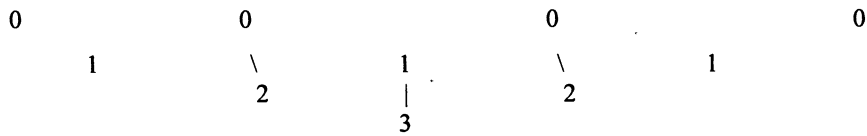


Figure 2. By shifting the right most “0” to the right, we will eventually reach this arrangement of roots.

4. GENERALIZATIONS. One might ask whether there is an obvious complex generalization to Corollary 2.2. If we have a polynomial with complex roots, and we move all the roots in one direction, will the roots of the derivative all follow? The answer, as expected, is no, as illustrated by the following counterexample: Take the third degree polynomial p which has complex roots i , $-i$, and a real root of value 2. One can check that the roots of p' are $\frac{1}{3}$ and 1. But if we shift to the right the real root of p from 2 to say 3, the roots of the derivative become $1 - \sqrt{\frac{2}{3}}$ and $1 + \sqrt{\frac{2}{3}}$. Since

$$1 - \sqrt{\frac{2}{3}} < \frac{1}{3} < 1 < 1 + \sqrt{\frac{2}{3}}$$

the two roots of the derivative did not *both* shift to the right.

REFERENCES

1. M. Marden, *Geometry of Polynomials*, American Mathematical Society, Providence, 1966.
 2. E. J. Barbeau, *Polynomials*, Springer-Verlag, New York, 1989.

*Department of Mathematics
 DePaul University
 Chicago, IL 60614
 anderson@math.cornell.edu*

Parallel Addition

Catherine C. McGeoch

If you set nine women to digging a ditch they will complete it in one-ninth the time required by a single woman. But nine women working together cannot bear a child in one month. The moral: some tasks can be parallelized and some cannot.

Can addition be parallelized? If one person can add two n-digit integers in n seconds, can n people add them in one second? It appears that n-digit addition requires n seconds no matter how many people are working on it, since the high-order digits cannot be added until the high-order carry-in is known. But in fact there does exist a method for adding integers in about 2 log2 n steps (using n people). The method is called carry-lookahead addition and is incorporated into the circuitry of nearly all modern computers. In this column we will look at carry-lookahead addition as well as an interesting parallel method for adding three integers.

We shall work with nonnegative n-digit integers expressed in binary (base two). Let X and Y be two such integers and let Z be their n + 1-digit sum. The digits of X are denoted xn-1xn-2...x1x0, and the digits of Y and Z are denoted similarly. Let C = cn-cn-1...c1c0 represent the carry digits: that is, ci is the carry-in added to xi and yi, and equivalently, the carry-out generated by adding xi-1, yi-1 and ci-1. We include c0 for notational convenience, recognizing that c0 = 0 always.

Figure 1-a contains a table defining one-digit binary addition with carry-ins and carry-outs. On the sixth line of the table, for example, we see that 1 + 0 + 1 = 10

Table with 5 columns: x, y, cin, cout, z. It shows binary addition results for combinations of x, y, and cin.

(a)

Table showing the addition of three numbers: T, C, X, Y, and Z. It includes a carry-in of 1 and shows the resulting sum Z.

(b)

Figure 1

in base two. Figure 1-b gives an example 6-digit sum, showing C, X, Y, Z and a row labeled T (described below). The usual way to add is to apply the one-digit function to the digits of X, Y , and C in turn as i goes from 0 to $n - 1$. The amount of time this takes (assuming constant time for each one-digit addition) is proportional to n . To achieve faster parallel addition we have to try something else.

Carry Lookahead Addition. Notice that we can sometimes calculate c_i without waiting to know the value of c_{i-1} . If x_{i-1} and y_{i-1} are both 0, then c_i must be 0, no matter what value c_{i-1} takes. Similarly, if x_{i-1} and y_{i-1} are both 1, then c_i must also be 1. The only problem arises when exactly one of x_{i-1} and y_{i-1} is 1, in which case c_i can't be determined until c_{i-1} is known.

We construct a *carry status* function f_i to reflect this situation. The carry status is expressed in terms of three functions k, g and p , each with domain $\{0, 1\}$. They are called the *kill* function, defined by $k(c) = 0$; the *generate* function $g(c) = 1$; and the *propagate* function $p(c) = c$. (You may recognize them by other names.) The carry status function is defined by

$$f_i(\cdot) = \begin{cases} k(\cdot) & \text{if } x_{i-1} = y_{i-1} = 0 \\ g(\cdot) & \text{if } x_{i-1} = y_{i-1} = 1 \\ p(\cdot) & \text{if } x_{i-1} \neq y_{i-1} \end{cases}$$

Row T in Figure 1-b shows the carry status functions for the example sum. It is easy to verify that $c_i = f_i(c_{i-1})$. Furthermore, we can apply function composition to obtain $c_i = f_i \circ f_{i-1}(c_{i-2})$. The handy table below shows the nine possible results of composing pairs of functions from $\{k, g, p\}$.

\circ	k	g	p
k	k	k	k
g	g	g	g
p	k	g	p

In general, $c_i = f_i \circ f_{i-1} \circ \cdots \circ f_j(c_{j-1})$ for $i > j > 0$. In particular, we have $c_i = f_i \circ \cdots \circ f_1(0)$, since by definition $c_0 = 0$. We will adopt the shorthand notation $[i, j]$ to refer to a sequence of compositions $f_i \circ \cdots \circ f_j$. In this notation $f_i = [i, i]$ and $c_i = [i, 1](0)$ for i between 1 and n . We stretch the notation slightly to let $c_0 = [0, 1](0) = 0$.

Carry-lookahead addition uses a clever two-pass scheme to find all the carries c_i quickly. In the first pass several compositions $[i, j]$ are calculated. In the second pass, functions of the form $[i, j](c_{j-1})$ are evaluated, one for each i between 1 and n . Once all the carry values $c_i = [i, j](c_{j-1})$ are known, the individual sums $x_i + y_i + c_i$ can be found simultaneously to produce the digits of Z .

Figure 2 shows a *combinational circuit* for performing carry-lookahead on 8-bit integers. The circuit comprises several *nodes* connected together by directed *wires*. The wires carry *values*: a node sends a value on its output wire according to some fixed function of values on its input wires. We require that each node have a fixed number of input wires and output wires, and that each node execute its function(s) in a fixed amount of time.

The circuit contains 7 *oval* nodes arranged in a binary tree, 1 *circle* node attached to the root of the tree, and 8 *square* nodes at the leaves of the tree. These different types of nodes perform different functions. In general, $n - 1$ ovals, 1

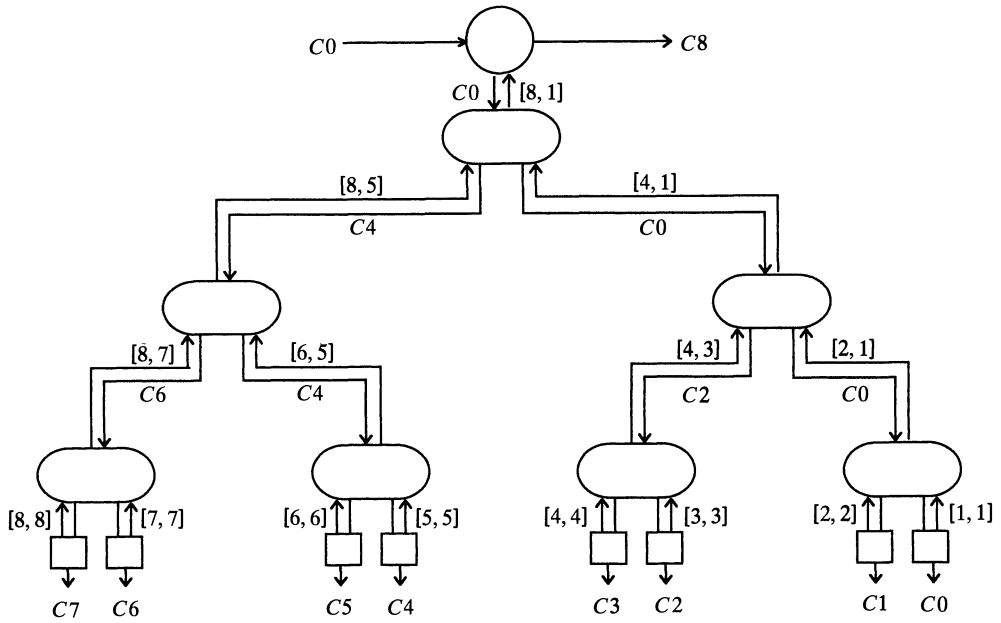


Figure 2

circle, n squares, and n more *adder nodes* (not shown here) are required for n -bit addition when n is a power of 2.

We can now add X and Y as follows.

Step 1. The n square nodes calculate $[i, i] = f_i(\cdot)$, for i in $1 \dots n$, each using inputs x_{i-1} and y_{i-1} (not shown in Figure 2). Each square node sends the appropriate value \mathbf{k} , \mathbf{g} , or \mathbf{p} along its output wire going up. This step requires $\Theta(1)$ time¹ since the square nodes can operate simultaneously.

Step 2. Each oval node performs the composition $[i, j] = [i, k] \circ [k - 1, j]$ *going up*. That is, the function values for $[i, k]$ and $[k - 1, j]$ (each is either \mathbf{k} , \mathbf{g} , or \mathbf{p}) are obtained from neighbor nodes below and the result $[i, j]$ is passed to the neighbor above. The circle node at top eventually receives $[n, 1]$. The arrows pointing up in Figure 2 are labelled to show the flow of values in this step. Overall, the time required for values to move from the square nodes (where the initial $[i, i]$ values are located) to the circle node (the last one to receive a value) is $\Theta(\log n)$.

Step 3. The circle node evaluates $[n, 1](0)$, equivalent to c_n . It also passes $c_0 = 0$ down to the root oval. Each oval node evaluates

$$c_{k-1} = [k - 1, j](c_{j-1})$$

going down. To accomplish this the node retains $[k - 1, j]$ from Step 2 and receives c_{j-1} from the neighbor above. The result c_{k-1} is passed down to the left neighbor, and $c_{j \perp 1}$ is passed down to the right. The arrows pointing down in Figure 2 are labeled to show the flow of values in this step. The total time required for values to propagate from the circle node down to the square nodes is $\Theta(\log n)$.

¹The notation $\Theta(f(n))$ means “proportional to $f(n)$ ” in the following sense: $g(n) = \Theta(f(n))$ means that there exist positive constants a and b and n_0 such that for all $n > n_0$ we have $af(n) \leq g(n) \leq bf(n)$. For example any constant function is $\Theta(1)$ and any function of the form $d \log_2 n + e$ (for constants d and e) is $\Theta(\log n)$.

Step 4. The circle node has computed c_n , and the square nodes now hold the carry values c_i for i between 0 and $n - 1$. These values can be passed to an array of adder nodes that simultaneously perform the 3-bit additions $x_i + y_i + c_i$ (discarding the carry-outs) to obtain the digits z_i of Z . This final step takes $\Theta(1)$ time.

The total amount of time required for the lookahead circuit and the adder array to form the sum of X and Y is $T(n) = L(n) + A(n)$, where $L(n) = \Theta(\log n)$ (to find the carry digits by lookahead) and $A(n) = \Theta(1)$ (to add the one-bit triples). Therefore $T(n) = \Theta(\log n)$. Note that although the circuit contains $3n$ nodes, carry-lookahead addition could be performed by n people acting as nodes, since people can move around and change functions.

Carry Save Addition. Now, how long does it take to add three n -digit integers W , X , and Y ? We could certainly add Y to the $n + 1$ -digit sum of W and X ; this would require $2T(n + 1)$ time if we use a lookahead-adder circuit of size $n + 1$ twice. A better idea is to apply *carry save* addition, which only requires $\Theta(1) + T(n + 1)$ time.

Given n -digit W , X , and Y , a carry save adder constructs two intermediate integers, an $n + 1$ digit U and an n -digit V , such that $U + V = W + X + Y$. Then U and V are summed with a carry-lookahead adder circuit of size $n + 1$.

Here's how it works. Referring again to Table 1-a, let the binary integer uv denote the 2-digit sum of three 1-digit binary integers; that is, $w + x + y = uv$. Then it must be the case that $w + x + y = 2u + v$.

For each i in $0 \dots n - 1$, apply the function in Table 1-a to the digits w_i , x_i , and y_i of W , X , and Y . Set $v_i = z$ and $u'_i = c_{out}$ as labelled in the table, and let v_i and u'_i denote the digits of V and U' , respectively. Let U be defined by $u_0 = 0$ and $u_i = u'_{i-1}$ for i in $1 \dots n$. Then V and $U = 2U'$ are the desired intermediate integers, since

$$\begin{aligned} W + X + Y &= \sum_{i=0}^{n-1} (w_i + x_i + y_i)2^i \\ &= \sum_{i=0}^{n-1} (2u'_i + v_i)2^i \\ &= \sum_{i=0}^{n-1} u'_i 2^{i+1} + v_i 2^i \\ &= \sum_{i=0}^{n-1} u_{i+1} 2^{i+1} + \sum_{i=0}^{n-1} v_i 2^i \\ &= U + V. \end{aligned}$$

The digits v_i and u_i can be calculated simultaneously by n adder nodes in $\Theta(1)$ time. After that, U and V can be added by a carry-lookahead circuit of size $n + 1$ in $T(n + 1)$ time.

Further Reading. Carry-lookahead addition and carry save addition have been around since the middle 1960's. We have since figured out how to parallelize several other arithmetic operations. For example, carry-save addition can be generalized so that a circuit containing $\Theta(nm)$ nodes can be used to add m n -digit numbers in $\Theta(\log_2 m + \log_2 n)$ time steps. This implies that two n -digit numbers can be *multiplied* in $\Theta(\log_2 n)$ time steps. We also know that under the standard

formal model of parallel computation it is *not* possible to add two n -digit numbers in $\Theta(1)$ time.

Two recent texts by Cormen et al. [1] and by Leighton [2] give excellent discussions of parallel arithmetic. The search for efficient parallel algorithms for general computational problems is a vigorous research area of theoretical computer science; Leighton's text, in particular, gives a comprehensive view of the state of the art.

ACKNOWLEDGMENT. This seems a good time to thank Dan Velleman for outstanding service as a "typical mathematical audience". Dan's insightful suggestions and comments on draft columns are most gratefully acknowledged.

REFERENCES

1. T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, MIT Press, 1990.
2. F. T. Leighton, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes*, Morgan Kaufmann, 1992.

Department of Mathematics and Computer Science
P.O. Box 2239
Amherst College
Amherst, MA 01002
ccm@cs.amherst.edu

Word has reached this country that the Editor of the *Zentralblatt für Mathematik und ihre Grenzgebiete*, Professor Otto Neugebauer, now of Copenhagen, has resigned. The resignation from this mathematical abstracts journal was occasioned by the action of the publisher, Julius Springer of Berlin, in dropping Professor Levi-Civita of Italy from the board without the knowledge of the Editor, as well as by the demand that the Editor give assurance that no emigrants would be allowed to referee articles by German authors. In consequence of this interference with editorial policies, the American associate editors, Professors Tamarkin and Veblen, have tendered their resignations as have also a number of associate editors and collaborators in other countries.

—*American Mathematical Monthly*
46 (1939), p. 57

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before April 30, 1994 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

An asterisk () after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10338. *Proposed by Charles Vanden Eynden, Illinois State University, Normal, IL.*

Given an integer $n > 1$, determine the set of integers which can be written as a sum of two integers relatively prime to n .

10339. *Proposed by Moshe Rosenfeld, Pacific Lutheran University, Tacoma, WA.*

Let A and B be complex matrices with $AB^2 - B^2A = B$. Prove that B is nilpotent.

10340. *Proposed by Richard Bagby, New Mexico State University, Las Cruces, NM.*

For a normed linear space \mathbf{X} and $x \in \mathbf{X}$, define

$$P(x) = \{y \in \mathbf{X}: \|x + y\|^2 = \|x\|^2 + \|y\|^2\}.$$

If the norm in \mathbf{X} comes from an inner product, then each $P(x)$ is invariant under multiplication by real numbers. Is the converse true?

10341. Proposed by George Cain and Zhiging Lu (student), Georgia Institute of Technology, Atlanta, GA.

Let $\mathbf{D} = \{(x, y): x^2 + y^2 \leq 1\}$ be the unit disk in the plane, and let $\{A_1, A_2, \dots, A_n\}$ be a pairwise disjoint collection of finite subsets of the set $\mathbf{C} = \{(x, y): x^2 + y^2 = 1\}$. Prove that there is a pairwise disjoint collection $\{K_1, K_2, \dots, K_n\}$ of connected subsets of \mathbf{D} such that $A_i \subset K_i$ for each $i = 1, 2, \dots, n$.

10342. Proposed by Shmuel Rosset, Tel Aviv University, Ramat Aviv, Israel.

Let F be a free group, and R a normal subgroup of F . Consider the subgroups $[R, nF]$ defined by

$$[R, nF] = \begin{cases} R & \text{if } n = 0, \\ [[R, (n-1)F], F] & \text{if } n > 0. \end{cases}$$

Prove that the set of elements of finite order in $R/[R, nF]$ is an abelian group.

10343. Proposed by David M. Bloom, Brooklyn College, CUNY, Brooklyn, NY.

Let us call a subset of \mathbb{Z} *semi-unfriendly* (abbreviated *S-U*) if it contains no three consecutive integers. Let E_n denote the n element set $\{1, 2, \dots, n\}$, and let

$$A(n, k) = \#\{S \subset E_n: \#S = k, S \text{ is } S-U\}$$

$$B(n, k) = \#\{S \subset E_n: \#S = k, S \text{ is } S-U \text{ and } E_n - S \text{ is } S-U\}.$$

Prove that

$$B(3n-1, n) = A(n+3, 3)$$

for all $n \geq 1$.

10344*. Proposed by E. Ehrhart, Université de Strasbourg, Strasbourg, France.

Let \mathcal{S} be a regular tetrahedron, and let $P \in \mathcal{S}$. Define $\mathbf{D}_V(P)$ to be the sum of the distances from P to the vertices of \mathcal{S} , and $\mathbf{D}_E(P)$ to be the sum of the distances from P to the edges of \mathcal{S} . Find the maximum and minimum values of $\mathbf{D}_E(P)/\mathbf{D}_V(P)$.

10345. Proposed by George Baloglou, SUNY College at Oswego, Oswego, NY, and Fred Galvin, University of Kansas, Lawrence, KS.

Given a subset $\mathbf{X} \subset \mathbb{R}$ one obtains a subset $\mathbb{R}^2 \setminus \mathbf{X}^2$ of the plane by removing those points both of whose coordinates are in \mathbf{X} . If $\mathbf{X} \neq \mathbb{R}$, such a set always contains horizontal and vertical lines.

(a) Find such a set \mathbf{X} , of Lebesgue measure zero, for which $\mathbb{R}^2 \setminus \mathbf{X}^2$ contains no circles.

(b)* Is there such a set \mathbf{X} , of Lebesgue measure zero, for which every connected subset of $\mathbb{R}^2 \setminus \mathbf{X}^2$ consisting of more than one point contains a horizontal or vertical line segment?

NOTES

Notes: (10339) An element, B , of a ring is called *nilpotent* if there is a positive integer k for which $B^k = 0$. For the ring of n by n matrices over the complex numbers, for fixed n , it would be of interest to seek a complete characterization of the solutions of the equation of this problem. **(10342)** Here, the symbol $[A, B]$ stands for the group generated by the commutators $aba^{-1}b^{-1}$ with $a \in A$ and $b \in B$. If A is a normal subgroup of B , so is $[A, B]$. A reference for free groups is Magnus, Karrass, & Solitar, *Combinatorial Group Theory*. **(10344)** This problem is listed as “unsolved” and no bounds are stated although partial results including conjectured extreme values are available, because these results are supported only by numerical evidence.

SOLUTIONS

Six Barycenters in Search of a Conic

E3469 [1991, 955]. *Proposed by Hüseyin Demir, Middle East Technical University, Ankara, Turkey.*

Suppose P is a point in the interior of triangle ABC and suppose AP, BP, CP meet the lines BC, CA, AB respectively at the points D, E, F . Prove that the centroids of the six triangles $PBD, PDC, PCE, PEA, PAF, PFB$ lie on a conic if and only if P lies on at least one of the three medians of the triangle.

Restatement of problem and fixing of notation. Applying the homothety with center P and ratio $3:2$ we see that the centroids of triangles are on a conic if and only if the midpoints of AF, FB, BD, DC, CE and EA are on one conic. Let x, y, z, u, v, w denote half the lengths of AF, FB, BD, DC, CE, EA , respectively. Let the midpoints of AF, FB, BD, DC, CE, EA be denoted by 1, 2, 4, 5, 6 respectively.

Solution 1 by Victor Prasolov, Independent University of Moscow, Moscow, Russia.

By Carnot's Theorem (see Howard W. Eves, *A survey of geometry* (Revised Edition), Allyn and Bacon, 1972, pages 256 and 262) the six centroids lie on a conic if and only if

$$x(2x + y)z(2z + u)v(2v + w) = w(2w + v)u(2u + z)y(2y + x). \quad (1)$$

By Ceva's Theorem, $xzv = wuy$, so (1) simplifies to $xzw + zvy + vxu - (wux + uyv + ywz) = 0$, or $(x - y)(z - u)(w - v) = 0$. This condition corresponds to P lying on a median.

Solution II by Albert Nijenhuis, Seattle, WA. By Pascal's Theorem, the points 1, 2, 3, 4, 5, and 6 lie on a conic if and only if the three points $Q = AB \cap 45$, $R = BC \cap 61$ and $S = CA \cap 23$ are collinear. (There is no real difficulty if any of these points are at infinity. The ratio AQ/QB , for example, is replaced by -1 if $AB \parallel 45$.)

By Menelaus' Theorem, we have

$$\frac{AQ}{QB} \cdot \frac{2z+u}{u} \cdot \frac{v}{2w+v} = -1, \quad \frac{BR}{RC} \cdot \frac{2v+w}{w} \cdot \frac{x}{2y+x} = -1,$$

$$\frac{CS}{SA} \cdot \frac{2x+y}{y} \cdot \frac{z}{2u+z} = -1.$$

Multiplying these together and using Ceva's theorem, as in Solution I, we see that $AQ/QB \cdot BR/RC \cdot CS/SA = -1$ if and only if $(x-y)(z-u)(w-v) = 0$. Thus Q, R, S are collinear and hence the points 1, 2, 3, 4, 5, 6 lie on a conic if and only if P is on a median.

Comments by Neela Lakshmanan, University of Scranton, Scranton, PA. The restriction that P is interior to the triangle may be relaxed: we need only that P does not lie on any side of the triangle.

We can prove that the result is true not only for the midpoints but also for the points that divide each of those six segments in a *constant ratio*: If 1, 2, 3, 4, 5, 6 are points on the sides of the triangle defined by $A1:1F = F2:2B = B3:3D = D4:4C = C5:5E = E6:6A$, then the six points lie on a conic if and only if P is on a median. Also, if P is an interior point, the hexagon 1, 2, 3, 4, 5, 6 is convex and attains its maximum area when P is the centroid of $\triangle ABC$.

Editorial comment. Many of the solvers supplemented the use of Carnot's Theorem or Pascal's Theorem with homogeneous coordinates and analytic methods. Some others worked directly with conditions on the six coefficients of a general conic.

Solved also by F. Bellot and M. A. Lopéz (Spain), R. J. Chapman (U.K.), J. Fukuta (Japan), H. Kappus (Switzerland), O. P. Lossers (The Netherlands), I. A. Sakmar (Turkey), Anchorage Math Solutions Group, and the proposer. One incorrect solution was received.

Periodicity of a Sign Function

E 3471 [1991, 955]. *Proposed by William Calbeck, Florida International University, Miami, FL, and Bruce Reznick, University of Illinois, Urbana, IL.*

Let P_k be the set of all integer-valued polynomials of degree at most k , i.e., the set of all polynomials p of degree at most k such that $p(n) \in \mathbb{Z}$ for $n \in \mathbb{Z}$. (It is known that $p \in P_k$ if and only if

$$p(x) = a_0 + a_1 \binom{x}{1} + a_2 \binom{x}{2} + \cdots + a_k \binom{x}{k},$$

where $a_0, a_1, a_2, \dots, a_k$ are integers.) Let $r(k)$ be the smallest power of 2 strictly greater than k .

(a) If $p \in P_k$, show that the sequence $\{(-1)^{p(n)}\}_{n=1}^{\infty}$ is periodic with period $r(k)$.

(b) Show that any given sequence of plus and minus ones with period 2^n occurs for some p in P_{2^n-1} .

Solution to part (a) by Robin J. Chapman, University of Exeter, UK. It suffices to show that if $j < 2^s$ and $m \in \mathbb{Z}$, then $\binom{m}{j} \equiv \binom{m+2^s}{j} \pmod{2}$. These are the coefficients of x^j in the power series $(1+x)^m$ and $(1+x)^{m+2^s}$, respectively. The congruence $(1+x)^{2^s} \equiv 1+x^{2^s} \pmod{2}$ follows easily by induction on s . Hence $(1+x)^{m+2^s} \equiv (1+x)^m(1+x^{2^s}) \pmod{2}$. Since j is less than 2^s , it is immediate that the coefficients of x^j in $(1+x)^m$ and $(1+x)^{m+2^s}$ have the same parity.

Solution to part (b) by Albert Nijenhuis, Seattle, WA. Let a_0, \dots, a_{2^n-1} be arbitrary integers, and let b_0, \dots, b_{2^n-1} be the solution to the equations

$$\sum_{i=0}^{2^n-1} b_i \binom{j}{i} = a_j \quad \text{for } 0 \leq j \leq 2^n - 1.$$

The matrix of this system is lower triangular, with 1's on the main diagonal, so $\{b_i\}$ are integers. The polynomial $\sum_{i=0}^{2^n-1} b_i \binom{x}{i}$ realizes the sequence $\{(-1)^{a_j}\}_{j=0}^{2^n-1}$ and its extension with period 2^n .

Editorial comment. Solvers used various methods; several cited theorems of Kummer and Lucas. William F. Trench recognized the problem as E 1365 [1959, 312; 1959, 919], proposed by M. E. Hausner and solved by N. J. Fine, and noted that a generalization to modulus m appears in William F. Trench, "On periodicities of certain sequences of residues," this MONTHLY 67 (1960), 652–656. Problem E 1365 had only two solvers in 1959.

Solved also by D. Callan, M. Dindos (Slovakia), F. J. Flanigan, R. High, K. S. Kedlaya (student), O. P. Lossers (The Netherlands), R. Martin (student), M. D. Meyerson, A. Pedersen (Denmark), W. F. Trench, Anchorage Math Solutions Group, National Security Agency Problems Group, and the proposer.

Arbitrarily Periodic Sequences

10184 [1992, 60]. *Proposed by Gerry Myerson, Macquarie University, New South Wales, Australia.*

Is there a sequence of natural numbers having the following two properties:

- (i) The sequence is periodic modulo m for every positive integer m ,
- (ii) each natural number appears in the sequence infinitely often?

Solution I by Kiran S. Kedlaya (student), Harvard University, Cambridge, MA. Yes. The following algorithm constructs such a sequence. Let S_1 be the sequence whose one term is 1. For $n \geq 1$, recursively define S_{n+1} by appending to the end of S_n the sequences $S_n + 0 \cdot n!$, $S_n + 1 \cdot n!$, \dots , $S_n + n \cdot n!$, where $T + k$ denotes the finite sequence obtained by adding k to each term of T . Let S be the sequence generated by this procedure as $n \rightarrow \infty$.

To prove that S satisfies (i), note that for all $n > m$, we obtain S_n by appending blocks that are congruent to S_m modulo m . Hence S is periodic modulo m with period dividing the length of S_m , which is seen by induction to be $(m+1)!/2$.

To verify that S satisfies (ii), first note by induction that S_n contains $\{1, \dots, n!\}$. Then observe that every number in S_n occurs at least twice as often in S_{n+1} . Thus every natural number appears in S infinitely often.

Solution II by Richard Stong, University of California, Los Angeles, CA. For each $n \geq 0$, there are unique integers c_1, \dots, c_k with $0 \leq c_j \leq j$ such that $n = \sum_{j=1}^k c_j j!$. Let $g(x) = \max\{0, x - 1\}$ and define $a_n = \sum_{j=1}^{k-1} g(c_{j+1})j!$.

Condition (i) holds for $\{a_n\}$ because $a_{n+(m+1)!} \equiv a_n \pmod{m!}$. To verify condition (ii), note that if $n = \sum_{j=1}^k c_j j!$, then $a_r = n$ for any r of the form $r = \sum_{j=1}^s b_j j!$ with $s > k$ and

$$b_j = \begin{cases} c_{j-1} + 1 & \text{if } 2 \leq j \leq k + 1 \\ 0 \text{ or } 1 & \text{otherwise} \end{cases}.$$

Editorial comment. Richard Stong's solution was the only submission giving an explicit formula for the n th term of the sequence; most solvers gave recursive procedures. Solvers disagreed on whether the natural numbers include 0. For this problem the question is moot, as the periodicity is unaffected by adjusting each term by 1. The proposer observed that "natural numbers" can be replaced by "integers" by alternating the terms of S with the terms of $1 - S$.

Solved also by M. Dasef & S. Kautz, P. Flor (Austria), J. Gonzalez-Meneses (student, Spain), J. W. Grossman, T. Hesterberg, R. High, N. Kang (student, Korea), U. Klein (student, Germany), O. P. Lossers (The Netherlands), M. D. Meyerson, A. Nijenhuis, I. Praton, A. Riese, R. M. Robinson, T. W. Starbird, D. M. Wells, GCHQ Problem Solving Group (U.K.), Theory First, University of South Alabama Problem Group, and the proposer.

A Golden Oldie

10193 [1992, 161]. *Proposed by Solomon Golomb, University of Southern California, Los Angeles, CA.*

Determine all pairs of integers n, k such that

$$\binom{n}{k} = \binom{n+1}{k-1}, \quad n > k > 1.$$

Solution by Christos Athanasiadis (student), Massachusetts Institute of Technology, Cambridge, MA. All such pairs are given by $n = F_{2m+1}F_{2m} - 1$, $k = F_{2m}F_{2m-1}$ for $m = 2, 3, \dots$. Here $\langle F_m \rangle_{m=1}^\infty$ is the Fibonacci sequence defined by $F_1 = F_2 = 1$ and $F_{m+2} = F_{m+1} + F_m$.

To see this, first note that the given condition can be written as

$$(n+1)k = (n-k+1)(n-k+2)$$

or as

$$(p+k)k = p(p+1), \tag{1}$$

where $p = n - k + 1$. Let $p = rt$, $k = st$ with $(r, s) = 1$. Then (1) becomes $(r+s)st^2 = rt(rt+1)$. It follows that t divides r , so that $r = tr_1$ and $(r+s)s = r_1(rt+1)$. Since r_1 is relatively prime to s and hence also to $r+s$, it must be that $r_1 = 1$. Hence $p = t^2$, $k = st$, and $t^2 + 1 = s(t+s)$. We need the following lemma.

Lemma. *The integer solutions to*

$$t^2 + 1 = s(t+s), \quad s \geq 1, t \geq 1 \tag{2}$$

are given by $s = F_{2m-1}$, $t = F_{2m}$ for $m = 1, 2, \dots$.

Proof: The classical formula $F_{2m+1}F_{2m-1} - F_{2m}^2 = 1$ (easily proved by induction) shows that $s = F_{2m-1}$, $t = F_{2m}$ is a solution of (2) for $m = 1, 2, \dots$. (In particular (1, 1) is a solution.) We now use an argument by descent to show that there are no other solutions of (2). Suppose that s and t are positive integers satisfying (2) and that $(s, t) \neq (1, 1)$. Put

$$\begin{pmatrix} v \\ u \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} t \\ s \end{pmatrix}, \quad \text{that is} \quad \begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v \\ u \end{pmatrix}. \quad (3)$$

Since $t^2 + 1 = st + s^2 > 1 + s^2$, we have $s < t$. Since $st + s^2 > t^2$, we have $(s/t) + (s/t)^2 > 1$ and hence $s/t > (\sqrt{5} - 1)/2 > \frac{1}{2}$. Thus $t/2 < s < t$, which implies that $u = 2s - t$ and $v = t - s$ are positive integers. It is easy to verify that $v^2 + 1 - u(v + u) = t^2 + 1 - s(t + s)$, so that (u, v) is a solution of (2) with $0 < u < s$, $0 < v < t$.

It follows by repetition of this argument that

$$\begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{m-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

for some positive integer m greater than 1. A simple induction argument shows that

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{m-1} = \begin{pmatrix} F_{2m-1} & F_{2m-2} \\ F_{2m-2} & F_{2m-3} \end{pmatrix} \quad (m = 2, 3, \dots).$$

Hence, if (s, t) is any solution of (2) other than (1, 1), we have

$$\begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} F_{2m-1} & F_{2m-2} \\ F_{2m-2} & F_{2m-3} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} F_{2m} \\ F_{2m-1} \end{pmatrix}$$

for some positive integer m greater than 1. Thus the lemma is proved.

In view of the lemma we have $k = st = F_{2m}F_{2m-1}$ and

$$n = p + k - 1 = t^2 + st - 1 = F_{2m}^2 + F_{2m}F_{2m-1} - 1 = F_{2m+1}F_{2m} - 1$$

for some integer m greater than 1, as claimed. The first five solutions (n, k) are (14, 6), (103, 40), (713, 273), (4894, 1870), and (33551, 12816).

Editorial comment. David M. Bloom and Savely Khosid each pointed out that the same problem appeared in the MONTHLY over sixty years ago as Problem 3459 [1930, 508; 1931, 551]. The above solution is more concise and direct than the solution published in 1931 (which used the theory of simple continued fractions). Problem 3459 is also the 65th problem in the collection of MONTHLY problems published as [1].

The problem is also treated in [3], [4], [5], and [6] (particularly pp. 32–34). These previous occurrences were called to our attention by B. M. M. de Weger, by Jean-Marie Pages and Dave Trautman, by Robert B. McNeill, and by Mark Sand respectively.

The diophantine equation $t^2 + 1 = s(t + s)$ of the above lemma may be written as $(2s + t)^2 - 5t^2 = 4$, an instance of the so-called Pell equation. (See, for example, Chapter 7 of Part One of [2].) Most solvers used the theory of the Pell equation or the theory of simple continued fractions. The selected solution bypasses the general theory, but uses knowledge of the small solutions of equation (2) to construct the change of variables in (3).

About one-third of the solvers obtained the result in the form $n = F_{2m}F_{2m+1} - 1$, $k = F_{2m}F_{2m-1}$, $m > 1$ given in the above solution. Several solvers included lists of pairs (n, k) produced by this formula. Some solvers, as well as reference [4], also gave the 29 digits of $\left(\frac{103}{40}\right)$. No one attempted to display the next value.

REFERENCES

1. The Otto Dunkel Memorial Problem Book, edited by Howard Eves and E. P. Starke, this MONTHLY, 64 (1957), no. 7, Part 2.
2. Edmund Landau, *Elementary Number Theory*, Chelsea Pub. Co., 1955.
3. D. A. Lind, "The quadratic field $Q(\sqrt{5})$ and a certain diophantine equation," *Fibonacci Quart.* 6, (1968), 86-93.
4. David Singmaster, "Repeated binomial coefficients and Fibonacci numbers," *Fibonacci Quart.* 13 (1975), 295-298.
5. Craig A. Tovey, "Multiple occurrences of binomial coefficients," *Fibonacci Quart.* 23 (1985), 356-358.
6. S. Vajda, *Fibonacci and Lucas Numbers, and the Golden Section*, Ellis Horwood Limited, 1989.

Solved by 88 readers and the proposer. Six incorrect solutions were also received.

Similar Orthic Triangles

10202 [1992, 265]. *Proposed by Juan Bosco Romero Márquez, Universidad de Valladolid, Valladolid, Spain.*

Let A', B', C' be the feet of the altitudes of $\triangle ABC$ and let X, Y, Z be the centers of the circumscribing rectangles of $\triangle ABC$ with edges BC, CA, AB respectively. Prove that $\triangle XYZ$ is a dilation of $\triangle A'B'C'$.

Solution I by Robin J. Chapman, University of Exeter, Exeter, U. K. There is an ambiguity as to what is meant by "circumscribing rectangle." The circumscribing rectangle of $\triangle ABC$ with edge BC may be defined as either:

- (i) the rectangle $BCPQ$ where A lies on the line PQ (possibly extended); or
- (ii) the smallest rectangle containing $\triangle ABC$, one of whose sides lies on the line BC .

These two definitions coincide provided neither $\angle ABC$ nor $\angle ACB$ is obtuse. The result is always true under interpretation (i), but false under interpretation (ii) whenever $\triangle ABC$ has an obtuse angle. In particular, if $\angle ABC$ is obtuse, then both X and Z coincide with the midpoint of AC , but $\triangle A'B'C'$ is not degenerate.

We adopt definition (i) and use vector methods. Choose the origin O to be the centroid of $\triangle ABC$. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{a}', \mathbf{b}', \mathbf{c}', \mathbf{x}, \mathbf{y}, \mathbf{z}$ be the position vectors of $A, B, C, S, A', B', C', X, Y, Z$ respectively. The circumscribing rectangle of $\triangle ABC$ with edge BC has vertices B, C and the points with position vectors $\mathbf{b} + (\mathbf{a} - \mathbf{a}')$ and $\mathbf{c} + (\mathbf{a} - \mathbf{a}')$. Hence $\mathbf{x} = (\mathbf{a} + \mathbf{b} + \mathbf{c} - \mathbf{a}')/2 = -\mathbf{a}'/2$ as O is the centroid of $\triangle ABC$. Similarly $\mathbf{y} = -\mathbf{b}'/2$ and $\mathbf{z} = -\mathbf{c}'/2$. Hence $\triangle XYZ$ is obtained from $\triangle A'B'C'$ by a dilation of factor $-1/2$ centered at the centroid O of $\triangle ABC$.

Solution II by Shailesh Shirali, Rishi Valley School, Chittoor District, Andhra Pradesh, India. Let $\triangle DEF$ be the medial triangle of $\triangle ABC$ with vertex D opposite vertex A . Then it is easy to see that $\triangle XYZ$ is just the orthic triangle of $\triangle DEF$ with vertex X opposite vertex D . Now, a dilation about the centroid G of $\triangle ABC$ with scale factor $-1/2$ sends $\triangle ABC$ to $\triangle DEF$ and therefore sends the orthic triangle of $\triangle ABC$, namely $\triangle A'B'C'$, to that of $\triangle DEF$, namely $\triangle XYZ$.

Editorial comment. Solution II, like other solutions employing constructions of classical geometry, was accompanied by a drawing. Jordi Dou submitted such a diagram entitled “Proof without words.” His diagram also highlights the fact, also observed by other solvers, that the dilation sending $\triangle XYZ$ to $\triangle A'B'C'$ also sends the circumcenter of $\triangle ABC$ (which is the orthocenter of the medial triangle) to its orthocenter, thereby exhibiting the fact that the centroid divides the segment joining the circumcenter and the orthocenter in the ratio of 1:2 (the property of the Euler line).

Jiro Fukuta proved the following more general result. Let A', B', C' be any points on the sides BC, CA, AB , respectively. Let X be the center of the circumscribing parallelogram with one edge BC and the other pair of edges parallel to AA' , and similarly for Y and Z . Then $\triangle XYZ$ is a dilation of $\triangle A'B'C'$, centered at the centroid of $\triangle ABC$, in the ratio of $-1/2$. This can be proved in the same way as the original problem, using either synthetic or vector methods. Since the lines AA', BB' , and CC' are not required to be concurrent, this is more general than the affine version of the stated problem.

This generalization can be easily carried over to higher dimensions in the following manner. Let $A_0 A_1 \dots A_n$ be a simplex in Euclidean n -space. For each $i = 0, 1, \dots, n$, let A'_i be any point in the facet opposite A_i , and let X_i be such that the vector $X_i - G_i$ is equal to the vector $(A_i - A'_i)/n$, where G_i is the centroid of the facet. Then $X_0 X_1 \dots X_n$ is a dilation of $A'_0 A'_1 \dots A'_n$, centered at the centroid of the given simplex, in the ratio $-1/n$.

Solved also by E. Alkan (student, Turkey), P. J. Anderson (Canada), J. Anglesio (France), F. Bellot and M. A. Lopéz (Spain), P.-C. Chuang, A. Coffman, I. Dimitric, J. Dou (Spain), J. Fukuta (Japan), H. W. Guggenheimer, J. G. Heuver (Canada), H. Kappus (Switzerland), I. Kastanas, K. S. Kedlaya (student), N. Komanda, O. P. Lossers (The Netherlands), M. Lucian, H. M. Marston, R. Merrill, K. Perera (student), W. Reyes (Chile), B. Shawyer (Canada), A. Subramanian (student, India), T. C. Tran, M. Vowe (Switzerland), R. L. Young, and the University of Wyoming Problem Circle. The original proposal presented only a special case of the published problem.

Matrices with Agreeable Adjoints

10205 [1992, 266]. *Proposed by Richard Sinkhorn, University of Houston, Houston, TX.*

In elementary linear algebra, two different definitions of the word “adjoint” are used. The adjoint of a square matrix A with complex entries is either:

- (I) the matrix whose (i, j) -entry is the cofactor of a_{ji} in A ; or,
- (II) the complex conjugate of the transpose of A .

Under what conditions on the matrix A will these two definitions yield the same matrix?

Solution by Peter Nylén, Tin-Yau Tam, and Frank Uhlig, Auburn University, Auburn, AL. There are three possibilities: (i) A is a zero matrix; (ii) A is unitary with $\det A = 1$; or (iii) A is a 2 by 2 matrix of the form

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}.$$

We use $\text{adj } A$ for the first “adjoint” and A^* for the second. The above matrices all satisfy $A^* = \text{adj } A$. Notice that $(\text{adj } A)A = A(\text{adj } A) = (\det A)I$. If $A^* =$

adj A , then

$$A^*A = AA^* = (\det A)I. \quad (1)$$

Since AA^* is positive semi-definite, $\det A \geq 0$. By taking the trace of both sides of (1), it follows that A is either nonsingular or zero. If A is nonsingular, take the determinant of both sides of (1). Then $|\det A|^2 = (\det A)^n$. Hence, if $n \neq 2$, $\det A = 1$, and consequently, $A^{-1} = (\det A)^{-1}(\text{adj } A) = A^*$, i.e., A is unitary. For $n = 2$, direct comparison of the entries of A^* and adj A gives the displayed form.

A related question is discussed in E. E. Underwood, "Classification of complex matrices A , where $A = \text{adj } A$," *Current Trends in Matrix Theory*, North-Holland, 1987, pp. 405–410. Michael K. Kinyon suggested replacing (1) by the "differentiated" condition

$$A + A^* = \text{tr}(A)I. \quad (2)$$

A similar method leads to the sequence of Lie algebras corresponding to the groups found above.

Solved also by D. Callan, R. J. Chapman (U.K.), I. Dimitric, W. T. Gan (student, U.K.), N.-G. Kang (student, Korea), M. K. Kinyon, N. Komanda, C. Lanski, F. Schmidt, R. Stong, E. T. Wong, University of Wyoming Problem Circle, and the proposer. Six incomplete solutions were also received.

Summing a Series of Volumes

10207 [1992, 266]. *Proposed by Eric Freden (student), Brigham Young University, Provo, UT.*

Find a closed form for $\sum_{n=0}^{\infty} \text{Vol}(B^n)$ where B^n is the unit ball in \mathbb{R}^n (and $\text{Vol}(B^0)$ is taken to be 1).

Composite solution by several solvers. More generally, if we take B^n as the ball of radius r in n dimensional space, then the series converges for all $r > 0$ and

$$\sum_{n=0}^{\infty} \text{Vol}(B^n) = e^{\pi r^2} \left(1 + \frac{2}{\sqrt{\pi}} \int_0^{r\sqrt{\pi}} e^{-t^2} dt \right).$$

This is proved in detail in D. J. Smith and M. K. Vamanamurthy, "How small is the unit ball?", *Math. Magazine* 62 (1989), 101–107.

Editorial comment. Most solvers indicated a reference both for $\text{Vol}(B^n)$ and the value of the resulting series. The terms of even dimension clearly determine an exponential function. The series consisting of the terms of odd degree can be recognized in terms of the solution of the initial value problem: $f'(x) = 1 + xf(x)$, $f(0) = 0$. Fourteen different references were given, none by more than three solvers.

Solved by K. F. Andersen (Canada), J. Anglesio (France), S.-J. Bang (Korea), W. H. Beckmann, D. M. Bloom, D. Callan, R. J. Chapman (U.K.), J. I. Concha (Chile), T. Dali and S. Smith and M. Carlton and P. Bracken, M. Dindos (Slovakia), M. Dresević and N. Cakić (Yugoslavia), M. Fichter (Germany), C. Georgiou (Greece), C. P. Grant, N.-G. Kang (student, Korea), M. K. Kinyon, N. Komanda, I. I. Kotlarski, R. Kreczner, O. P. Lossers (The Netherlands), S. Matz, A. Pedersen (Denmark), K. Perera (student), F. C. Rembis, R. M. Robinson, P. Sawyer (Canada), B. D. Sterba-Boatwright, R. S. Tiberio, A. Tissier (France), D. B. Tyler, D. C. Vella, M. Vowe (Switzerland), D. M. Wells, P. J. Zweir, National Security Agency Problems Group, Shreveport Problem Solving Group (LSU), University of Wyoming Problem Circle, and the proposer. One incorrect solution was received.

Collaborating editors: David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Jerrold R. Griggs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pfeifer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, and William E. Watkins.

Answer to Picture Puzzle

(p. 847)

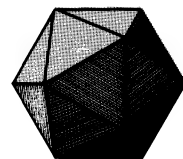
Louis de Branges, the solver of the Bieberbach conjecture.

During the last quarter of a century there has been a universal effort to improve the quality of teaching in the elementary and secondary schools. Whenever a change is made in this country in the curricula for the training of teachers, it has been in the direction of more "education", pedagogy and psychology, always at the expense of further courses in subject matter. The results are already apparent; for grade schools the new method is an improvement, but for high schools, especially the last two years, it is lamentably deficient. However desirable the other things may be in themselves, for a teacher of mathematics nothing has yet been discovered to replace a knowledge of mathematics. May the present volume take its place in American and English schools, to extend the service it has so admirably rendered in Germany.

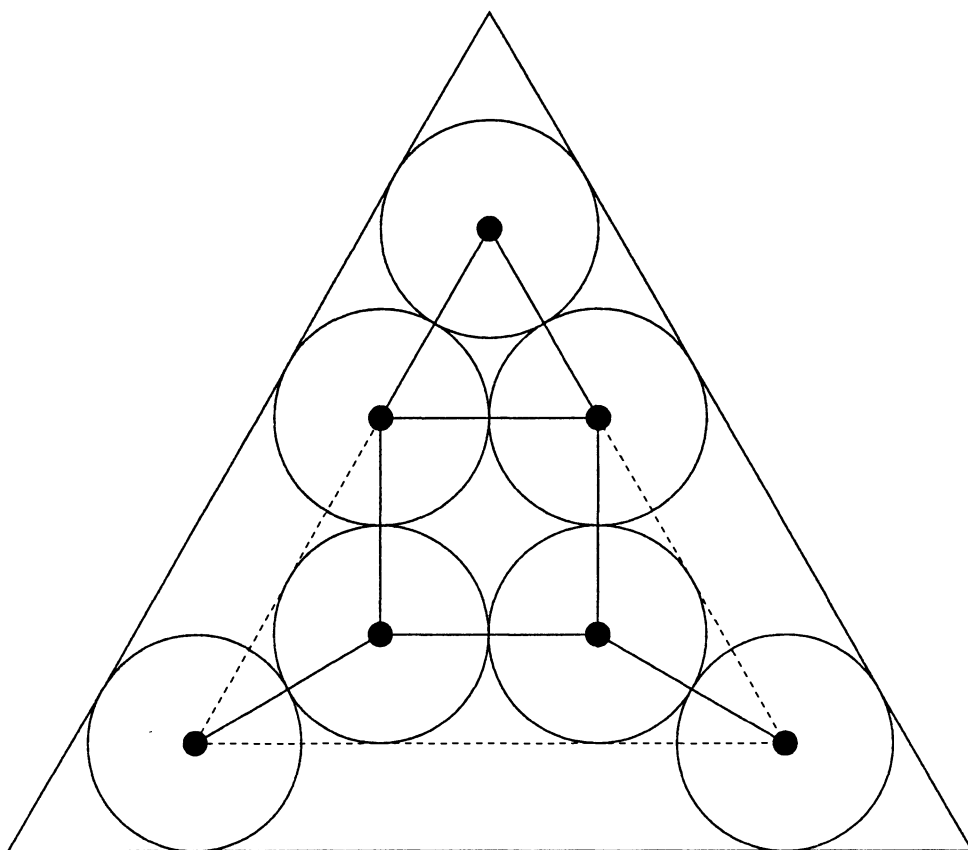
VIRGIL SNYDER

—*American Mathematical Monthly*
40 (1933), p. 171

The American Mathematical Monthly



Volume 100, Number 10 / DECEMBER 1993



NOTICE TO AUTHORS

The *Monthly* publishes articles, notes, and other features about mathematics and the profession. The readership of the *Monthly* is intended to include everybody who is mathematically inclined, including of course professional mathematicians and students of mathematics at all collegiate levels. While no single article or feature is likely to appeal to everyone, material should interest and be accessible to a large number of readers. This is the most important criterion for acceptance.

Articles may be expositions of old results or presentations of new ones. They may concern all of mathematics or one small area, a broad development or a single application, historical reminiscences or one important event. While some articles may contain the author's new research, the novelty of material and generality of the results is far less important than the clarity of exposition and general interest. Discussing one illuminating case of a well known result is far better than providing all the details of an obscure but new proposition. Articles in the *Monthly* are supposed to inform and to entertain; they are meant to be read rather than archived.

Notes are short and possibly informal articles. A note may concern a clever new proof of an old theorem, a novel way to present tired material, or a lively discussion of a philosophical (but still mathematical) issue. Also, any topic is suitable, so long as it is related to mathematics. Because a note is short, the first few sentences are the most important part: They should explain the purpose and invite the reader in. Photographs or diagrams often will attract the reader's attention.

All articles and notes should be sent to the editor:

JOHN EWING
Department of Mathematics
Indiana University
Bloomington, IN 47405

Please send 3 copies, typewritten on only one side of the paper. Illustrations should be carefully drawn on separate sheets of paper in black ink; the original should be without lettering and two copies should have appropriate captions and lettering indicated.

Proposed problems or solutions should be sent to:

RICHARD BUMBY,
P.O. Box 10971
New Brunswick, NJ 08906-0971.

Please send 2 copies of all material, typewritten if possible.

Letters to the Editor, both for publication and for private reading, should be sent to the Editor at the address given above. Comments, including criticisms, are welcome, as are all suggestions for making the *Monthly* a lively, entertaining, and informative journal.

EDITOR:

JOHN H. EWING

ASSOCIATE EDITORS:

RONALD BOOK	JOAN HUTCHINSON
PETER BORWEIN	CATHERINE MCGEOCH
RICHARD BUMBY	RICHARD NOWAKOWSKI
DENNIS DETURCK	ARNOLD OSTEBEE
UNDERWOOD DUDLEY	LEE RUBEL
JOHN DUNCAN	LYNN STEEN
JOAN FERRINI-MUNDY	STAN WAGON
JOSEPH GALLIAN	DOUGLAS WEST
STEVEN GALOVICH	HERBERT WILF
RICHARD GUY	SANDY ZABELL
DARRELL HAILE	PAUL ZORN
PAUL HALMOS	

EDITORIAL ASSISTANT:

MISTY CUMMINGS

STAFF ARTIST:

MIKE CAGLE

Reprint permission:

MARCIA P. SWARD, Executive Director

Advertising Correspondence:

Ms. ELAINE PEDREIRA, Advertising Manager

Subscription correspondence, change of address, and other inquiries:

Membership / Subscriptions Department

All at the address:

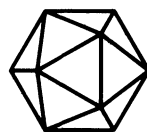
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036.

Microfilm Editions: University Microfilms International, Serial Bid coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

The AMERICAN MATHEMATICAL MONTHLY (ISSN 0002-9890) is published monthly except bimonthly June-July and August-September by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, DC 20036 and Montpelier, VT. Copyrighted by the Mathematical Association of America (Incorporated), 1994, including rights to this journal issue as a whole and, except where otherwise noted, rights to each individual contribution. General permission is granted to Institutional Members of the MAA for noncommercial reproduction in limited quantities of individual articles (in whole or in part) provided a complete reference is made to the source. Second class postage paid at Washington, DC, and additional mailing offices. **Postmaster:** Send address changes to the American Mathematical Monthly, Membership / Subscription Department, MAA, 1529 Eighteenth Street, N.W., Washington, DC, 20036-1385.

**The American
Mathematical Monthly**

Volume 100 Number 10 / DECEMBER 1993
(ISSN 0002-9890)



Contents

ARTICLES

Thomas Archer Hirst—Mathematician Xtravagant VI. Years of Decline /
J. HELEN GARDNER and ROBIN J. WILSON 907

Densest Packings of Congruent Circles in an Equilateral Triangle /
HANS (J. B. M.) MELISSEN 916

Partnerships / ALAN H. SCHOENFELD 926

A Simple Proof of Pascal's Hexagon Theorem / JAN VAN YZEREN 930

FEATURES

COMMENTS 906

NOTES 932

POSTCARDS FROM MAX / PAUL HALMOS 942

UNSOLVED PROBLEMS

A Quarter Century of *Monthly* Unsolved Problems, 1969–1993 /
RICHARD K. GUY 945

THE AUTHORS 950

PROBLEMS AND SOLUTIONS 951

REVIEWS

A First Course in Chaotic Dynamical Systems: Theory and Experiment.
By Robert Devaney. *A First Course in Chaotic Dynamical Systems:*
Labs 1–6. By James Georges, Del Johnson and Robert Devaney /
PHILIP D. STRAFFIN 961

TELEGRAPHIC REVIEWS 964

INDEX TO VOLUME 100 970

Thomas Archer Hirst— Mathematician Xtravagant VI. Years of Decline

J. Helen Gardner and Robin J. Wilson

I have had several letters during the week from Cayley on Geometrical Transformation. I wish I were at liberty to do my part in the important investigations that are now ripe; but I have to exercise self-denial. My lectures absorb my time and constitute my duty. Sylvester again is actively thinking and producing, and Chasles has just published a most important extension of his method. I must simply look on.

By 1865, Thomas Hirst was at the height of his powers. As Professor of Mathematical Physics at University College, London, Vice-President of the newly-formed London Mathematical Society, a member of the distinguished X-club, and a Council member of the Royal Society, he was in a position to influence those around him. A long-standing ambition was to propose the French geometer Michel Chasles for the Royal Society's Copley medal.

29th October 1865: ... my proposition (although late) was well received; it was unanimously agreed that his name should be put on the list. The adjudication is on Thursday next, and I shall work hard to carry him. He has formidable rivals however in Regnault, Plücker, and Poncelet.

And he was successful, although the ailing and elderly Chasles was too unwell to come to London for the ceremony. At the celebration dinner afterwards, a toast was proposed to Chasles, the Copley Medallist and 'his friend Dr Hirst'. Following this toast, Hirst made a speech describing Chasles's achievements, which he included in full in his diary entry. He was obviously very pleased with his success, and now looked forward to presenting Chasles with his prize.

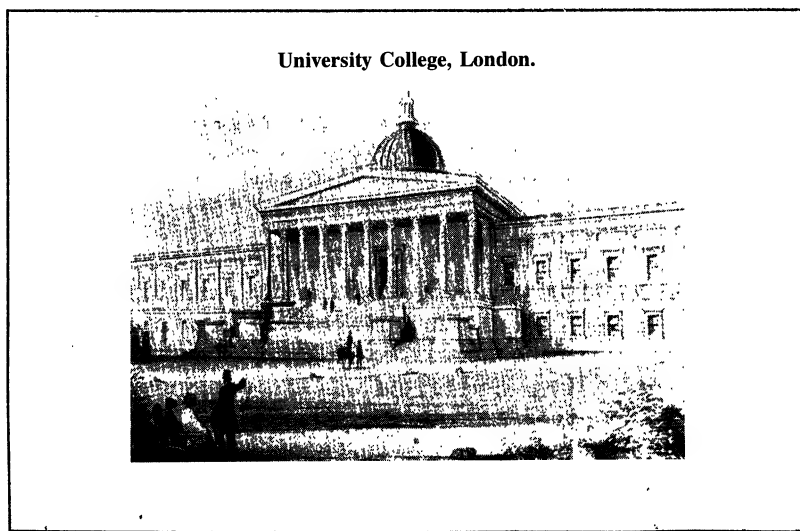
30th November 1865: ... I have but one step more to take and that will be across the channel to the Passage St. Marie, Rue de Bac at Paris, there with my own hands I will place the medal in the hands of Chasles, as a grateful offering to the man who, next to Steiner, has been most influential in determining my own career.

24th December 1865: ... My first act this morning was to call on Chasles and deliver the Copley Medal. It was manifestly a welcome present to him...

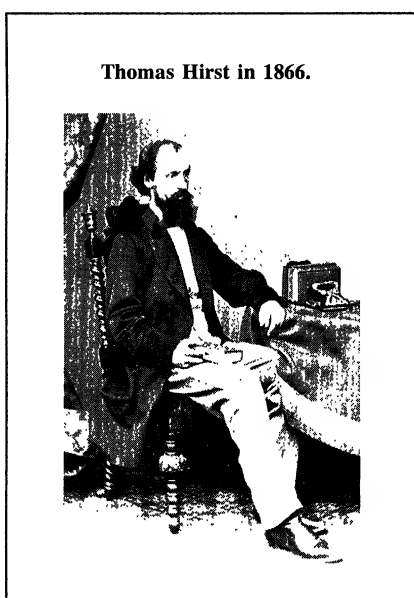
Throughout 1866, Hirst added further to his list of personal achievements. In February, he was elected a Member of the Athenaeum Club, in June he was appointed General Secretary of the British Association for the Advancement of Science, an onerous post which he held for four years, and in November he was admitted a Fellow of the Royal Astronomical Society. While attending a rather uninteresting meeting of one of his clubs, he 'made a calculation to show that there would be ample standing room for all the inhabitants of the Globe in the Isle of Wight'—a result which greatly surprised him.

Meanwhile, the London Mathematical Society was quickly becoming established. At one meeting, the President, Augustus De Morgan, 'called attention to the novelty and importance of many of the papers, and remarked that this was the only society in England where such papers could be received'. It seems that the Society was keen to encourage young talent:

22nd November 1866: At Math. Society. Clifford of Trin. Coll. Cambridge made his first appearance and gave us a very good paper 'on Harmonics'. There is no young mathematician of greater promise than Clifford just now.



In 1867 Augustus De Morgan had a disagreement with University College, and resigned his Chair in Mathematics. Hirst was elected in his place 'unconditionally and most unanimously'. He proved to be a first-class choice, if the memory of one of his students is accurate.



'His presence in the classroom was striking. He was tall, and held himself erect with an almost military air. He had a long black beard and a great, bald, dome-like forehead. He was a man with whom it was impossible to imagine the most audacious student venturing to take a liberty. There was something about him that invested his unlovely subject with dignity, if not interest. Less, perhaps, than any of the other professors, did he seem to think of examinations. To him, I believe, incredible as it sounds, mathematics must have been a solemn, high pursuit: a passion, if not a religion. Yet with all his aloofness of manner he could be very simple, very patient, and extremely kind. Certainly to one of his most hopeless pupils he showed himself all three.'

Meanwhile, the X-club continued its tradition of monthly meetings, with the occasional distinguished guest in attendance.

3rd March 1868: At the X-Club. Darwin was our guest. I was in the chair, and again the evening passed very pleasantly away.

2nd April 1868: At the X. Huxley, Frankland, Sir J. Lubbock and myself were the only ones who dined. Spottiswoode was there for an hour and brought Clifford with him. Clifford is the Lion of this season. Everybody is anxious to entertain him. I hope only his head will remain unturned.

But increasingly he found that his administrative and lecturing duties left him too little time for his researches, and he frequently complains of his inability to spend enough time on geometry.

7th February 1869: At home writing paper on Degenerate Conics. This paper perplexes me sorely, I begin to fear that it will never be satisfactorily written until I can work at it uninterruptedly. My daily duties so absorb my thoughts that I can only in leisure hours succeed in turning them to this new work, and no sooner are they turned and effective work rendered possible than the said duties turn them away again.

Tyndall, generous friend, proposed a remedy for this incessant disappointment I experience which I must record; it was so characteristic. "Give up your Professorship and devote yourself for a few years to your work solely. I have more money than I want and I can easily spare you what you would require to enable you to work without embarrassment."

However dear to me the privilege of thus working I could not, of course, accept it on these easy terms. My first duty is to earn my bread by teaching; if original research is not compatible with the performance of this duty then I must sacrifice originality however dear to me it may be, or however much my science might be advanced thereby. If the mathematical world prefer my teaching to my researches what right have I to complain? Can I even say that its choice is a bad one? I doubt it.

Tyndall realized that Hirst's researches could lead to important discoveries. Indeed, had he managed to persuade Hirst to take up his offer, Hirst's name might have been better remembered. As it was, the situation did not improve, and two days after his 39th birthday, Hirst wrote:

24th April 1869: Working at quadric transformation. Cayley and Clifford have begun to work at the subject and unless I communicate what I did in 1865 I shall be out-run. How I long to have leisure to pursue my work. So long as my present drudging continues I shall be scientifically speaking extinguished.

Despite this, or perhaps because of a need for relaxation from the pressure he was under, Hirst made one of his regular visits to the Continent to meet old and new acquaintances:

26th July 1869: Bath in Neckar. We walked up to the Castle and saw all over it, the Fass included. Dined at Hotel Schrieder at 1. P.M. with Bunsen where we met Kirchhoff (on crutches) and Königsberger the Mathematician and successor of Hesse, now at Munich. We took our Abendessen with Helmholtz...

27th July 1869: After another bath in the Neckar I attended Königsberger's lecture on Theory of Determinants. He introduced me to a young Russian lady [Sonya Kowalevskaya]... who attends his lectures and is at home in Elliptic Functions. She belongs to the mathematically gifted family of Schuberts. She is pretty and exceedingly modest.

Back in England, Thomas Hirst's teaching activities took a new direction:

Ladies' Educational Association, London. A Course of Twenty-four Lectures on the Elements of Geometry will be given by Professor Hirst, in the Minor Hall, St. George's Hall, Langham Place, on Mondays and Fridays at 11. A.M. (beginning on January 17), should a sufficient number of tickets be applied for before Christmas. The Lectures will be of an elementary character requiring no previous knowledge of the subject, the extent to which it will ultimately be carried being dependent upon the progress of the class.

Fee for the Course of 24 Lectures, £11.1.6; Governesses £1.1s. Ladies over seventeen years of age may join this or any other Course in connection with the Association (that of Chemistry subject to the approval of the lecturing professor) after Christmas on the above reduced terms.

Thirty ladies enrolled for the first lecture, but about sixty attended. By the following week, fifty-seven students had enrolled, and a measure of Hirst's exceptional teaching skills may be gained in that half-way through the course he records that "one or two only have confessed inability to follow".

After long deliberation, he made up his mind to resign his chair, and apply for a well-paid administrative job which he hoped would give him more time for his research.

28th February 1870: ... The fact that I cannot at present do any original work, that it is only by devoting myself wholly to lecturing that I can keep up my number of students at the College and thus secure my bread; that as my strength fails my prospects will necessarily be worse at University College; these facts I say decided me at length to apply for an appointment of an inferior order, perhaps, but of a less arduous and more remunerative character. Moreover if I succeed I shall come in contact with good and influential men and myself be able to influence to some extent the character of Education in England.

After some confusion, in March 1870 the Senate of the University of London appointed him Assistant Registrar and, for a time, his researches began to make progress again. He began work on a memoir on the "Correlation of two Planes".

31st December 1871: ... It grows under my hands both in bulk and, I think, in value. Small as is my year's achievement, it has given to my life a purpose for which I feel grateful. It has raised my life in my own estimation,—and it is almost the only thing that has done so—above mere routine and mediocrity. To keep my brain clear and in a condition to discover geometrical relations has become to me a main purpose in life, all other objects have in comparison become of little moment to me.

Hirst now found time to devote to a topic which had been dear to his heart for several years. Already by 1868 he had come to believe that Euclid's *Elements* should be supplanted as the main geometry textbook in English schools, and accordingly he had spent some time editing a new geometry book by Richard Wright. This conviction, arising from his years as a surveyor and his experience of teaching practical geometry at Queenwood and University College School, left him well placed to help establish a new association whose aim was to reform the teaching of geometry in schools. This was the Association for the Improvement of Geometrical Teaching, which was founded in January 1871; Hirst was its first president, and held office for seven years. Later, in the 1880s, it broadened its scope to cover the whole range of school mathematics, and in 1897 it was re-named the Mathematical Association, a name which it holds to this day.

In 1872, Hirst was elected President of the London Mathematical Society for a period of two years. Sylvester had suggested Cayley for this post, and Hirst was also proposed, despite wanting to remain Treasurer.

The Royal Naval College in Greenwich.



10th October 1872: ...At the first vote Cayley stood first, I next and Henrici last but none obtained an absolute majority of votes. Henrici's name was accordingly withdrawn and the voting resumed when I obtained one more vote than Cayley. I voted for Cayley both times... Had Spottiswoode not strongly urged my accepting the office of President and had it been any other than Sylvester who divided the Council between Cayley and myself I should have persisted in declining to serve in any other capacity than that of Treasurer. Sylvester's *animus* against me was disagreeably manifest. It has lasted now for years and the cause of it is just as unknown to me as it was on its first appearance. ...

In the following year, Hirst embarked on his fourth (and final) career. He was appointed the first Director of Studies at the Royal Naval College in Greenwich, with a salary of £1200 per year, plus a house. This position enabled him to keep in touch with the international mathematical community.

3rd October 1873: Tchebichef, who called on me a few days ago, and Klein dined with me at Greenwich. Tchebichef told us of a mode of converting circular into rectilineal motion (à propos of the parallelogram of Watt) which was a simple and beautiful application of Quadric Inversion.

For some years there had been a lack of contact between Hirst and Sylvester. In 1875, on learning that the latter was suffering from rheumatism in the eyes, Hirst broke the long silence by expressing his sorrow at Sylvester's affliction.

25th May 1875: ...He voluntarily shook hands with me, and thus at last there is a kind of reconciliation between us. I am very glad of it, though I have learned to my sorrow that our former intimacy can never be renewed. What the exact cause of our original estrangement was I never knew, but I do know that he suspected me most unjustly of incessantly plotting to undermine his influence in the scientific and mathematical circles. He misconstrued every act and word of mine to such an extent that intercourse was impossible.

In 1878, his work was recognized by the University of Cambridge:

8th June 1878: I received the Diploma of Membership of the Cambridge Philosophical Society. I was gratified about a month ago to hear through Glaisher of my election. I may say that this is the first recognition I have ever received from any University in my native country.

The next year, he made yet another visit to the Continent. While in Paris, he met Liouville in the street.

18th May 1879: ... A little shrivelled gouty old man he has become and very garrulous. It was with difficulty I broke away from him ...

More enjoyable was a visit to Parpan in Switzerland, a little Alpine village 4,545 feet above the sea, where 'I found Cremona, Casorati, Beltrami (with the Signora C), Geiser, Schlöffli, Frobenius and Meier (seven mathematicians!)'.

Then, in late 1880, he learned that his unpublished researches had indeed been out-run, as he had forecast in 1869.

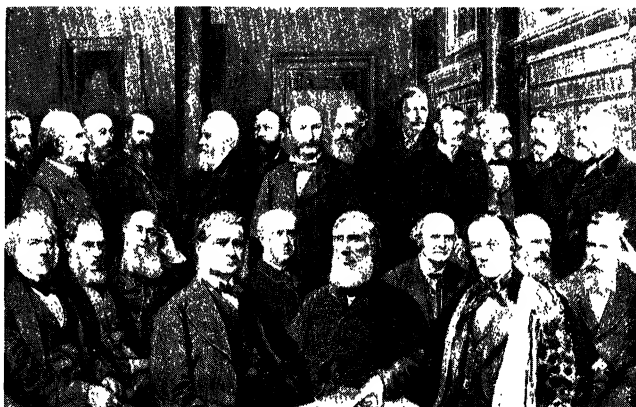
11th November 1880: The first meeting of the Math. Soc. took place on Nov. 11th. Cayley came to it and stopped with me. We were speaking of Cantor's paper on the cyclical self-corresponding points in two coincident planes between which a quadric relation exists. It has just appeared in the *Annale di Matematica*. I communicated precisely the same theorem to the British Association at Birmingham in 1865 but nothing was printed about it except the barest notice in the Proceedings. I showed Cayley my M.S. notes for that communication. He took them home with him and expressed an intention to write something about the matter. I shall be glad to be associated with a theorem which was always a pet of mine. As usual however I went on nursing my pet with the intention of allowing it to grow and develop itself more before I published it.

In 1883 he heard from Thomas Huxley that the Royal Society had awarded him its prestigious Royal Medal, principally for his work on Cremona transformations:

30th November 1883: ... I received my Royal Medal from Huxley who addressed to me a few friendly words in addition to the formal ones of presentation. "Although quite out of order" he said "I cannot refrain from expressing my sincere pleasure at being able, on the first occasion of my official representation of the Royal Society, to hand this Royal Medal to one of my oldest friends".

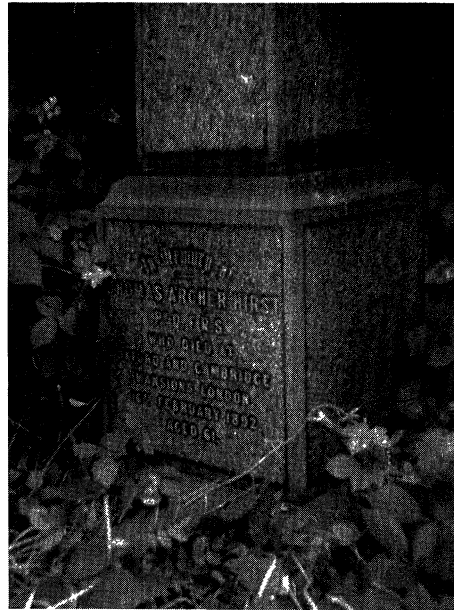
The Royal Society—a portrait group of some of the most distinguished Fellows in 1889.

At the front are, from left to right, Sir Gabriel Stokes, Sir Joseph Hooker, James Joseph Sylvester, Thomas Huxley, Archibald Geikie, John Tyndall, Arthur Cayley, Sir Richard Owen, W. H. Flower, and William Crookes.



The last entries of Hirst's diary, and Hirst's grave in Highgate Cemetery, North London.

...the last entries of Hirst's diary, and Hirst's grave in Highgate Cemetery, North London.



Hirst's health had always been a cause of concern, and now it continued to decline as first kidney stones and then a stomach tumour were diagnosed. He found life very lonely when his brother John died and his favourite niece, who at one stage had been his housekeeper, married and then died in childbirth. He travelled to Greece and Egypt and in 1883 he gave up his Greenwich post at the age of 53. He now had the time to work on his geometry, at his clubs in London during the summer, and in France during the winter. He also featured in a popular book:

11th January 1890: I gave some final touches today to the notice of myself and my work in "Men of the Time". It will be posted tomorrow.

Finally, in 1890, he finished his memoir on the correlation of two spaces. He had worked on it for a long time, and after its completion he destroyed his mathematical notebooks. Suddenly he seemed old, spending his time in watching the rapidly changing world from his clubs, his flat, and the park:

23rd August 1890: ... What a mad world it is! In the distance the Sunday Band was playing unmelodiously. What a noisy, jiggling world it has become!

He became increasingly depressed by the number of his colleagues and acquaintances who were departing this world.

19th February 1891: ... At the Athenaeum I read, in Nature, of the death of Madame Sophie Kovalevsky (aged 38), Professor of Mathematics at the Högskola of Stockholm. When she was 18 years of age I was introduced to her by Königsberger at Heidelberg, whose lectures she was then attending. Some years afterwards she studied under Weierstrass at Berlin... As far as her

mathematical abilities were concerned, she appears to have been superior to any predecessor of her own sex. She died from an attack of pleurisy; brought on, it is believed, by a chill which succeeded her rapid journey home from the South of France in order to commence her lectures at Stockholm.

For thirty-four years Anna had never been far from his thoughts and in September 1891, he paid one of his regular visits to Paris to bid Anna good-bye for the last time. The turn of the year brought yet more sad news.

7th January 1892: I hear from Sturm this morning that Heinrich Schröter, of Breslau, is dead. He and I heard Steiner's lectures together, at Berlin, in 1851–2... He has been taken before me. When will my time come?...

It came sooner than he thought. London was hit by a flu epidemic, one of the worst of the century. His resistance lowered by years of illness, and now suffering from cancer of the prostate, Hirst quickly succumbed. His last diary entries were written just four weeks before his death.

17th January 1892: ... the symptoms of violent cold in the head continued until nearly midnight. I then went to bed, but slept only in a disturbed fashion and awoke with pains and cramps all over my body. I fear the influenza has overtaken me.

18th January 1892: I rose in a sad plight. I took coffee for breakfast, however. This set the bowels acting; but no relief from my oppressive malaise followed. Cranstone called to look at the fallen chimney-piece in my sitting room.

On 16th February 1892 he died, and was buried in Highgate Cemetery.

Thomas Archer Hirst—Principal publications

1. On the existence of a magnetic medium, *Proc. Roy. Soc.* 7 (1854/5), 448–454.
2. On equally attracting bodies, *Phil. Mag.* 13 (1857), 305–324.
3. On equally attracting surfaces, *Phil. Mag.* 16 (1858), 160–177, 266–284.
4. On derived surfaces, *Quart. J. Pure Appl. Math.* 3 (1860), 210–218.
5. On ripples and their relation to the velocities of currents, *Phil. Mag.* 21, (1861), 188–198.
6. On the volumes of pedal surfaces, *Phil. Trans.* 153 (1863), 13–32.
7. On the quadric inversion of plane curves, *Proc. Roy. Soc.* 14 (1865), 91–106.
8. On the degenerate forms of conics, *Proc. London Math. Soc.* 2 (1866/9), 166–173.
9. On the correlation of two planes, *Proc. London Math. Soc.* 5 (1873/74), 40–70.
10. On correlation in space, *Proc. London Math. Soc.* 6 (1874/5), 7–9.
11. Notes on the correlation of two planes, *Proc. London Math. Soc.* 8 (1876/7), 262–273.
12. Note on the complexes generated by two correlative planes, *Proc. London Math. Soc.* 10 (1878/9), 131–153.
13. On quadric transformation, *Quart. J. Math.* 17 (1881), 301–311.
14. On Cremonian congruences, *Proc. London Math. Soc.* 14 (1882/3), 259–301.
15. On congruences of the third order and class, *Proc. London Math. Soc.* 16 (1884/5), 232–237.
16. On the Cremonian congruences which are contained in a linear complex, *Proc. London Math. Soc.* 17 (1885/6), 287–296.
17. Translation of R. J. E. Clausius, *The mechanical theory of heat with its applications to the steam engine and to the physical properties of bodies*, London, 1887.
18. On the correlation of two spaces, each of three dimensions, *Proc. London Math. Soc.* 21 (1889/90), 92–118.

ACKNOWLEDGMENTS. A typed version of the Thomas Hirst diaries is held at the Royal Institution in London, and quotations from the diaries appear here by courtesy of the Royal Institution. The diaries have been edited by W. H. Brock and R. M. MacLeod, and were published in microfiche by Mansell, London, in 1980.

PICTURE CREDITS. Maps of Yorkshire and Germany and Hirst's grave, courtesy Helen Gardner; Halifax (from a 19th-century print of J. R. Smith); De Morgan, Cremona, and Hirst (profile), courtesy The London Mathematical Society; Tyndall lecturing (from *The Illustrated London News*, 14th May 1870), and the Royal Society (from *The Illustrated London News*, 12th December 1863), courtesy The Illustrated London News Picture Library; Faraday (from a 19th-century print of McGuire), courtesy The Royal Institution; Hirst (portrait), Royal Naval College, Greenwich; Marburg, the University of Marburg, and Hirst's dissertation, courtesy Picture Archive, Philipps University of Marburg, Germany; Bunsen (engraved from a 19th-century photograph by C. Cook); Göttingen (from H.-H. Himme, *Stich-haltige Beiträge zur Geschichte der Georgia Augusta in Göttingen*, Vandenhoeck und Ruprecht, 1987), courtesy Vandenhoeck und Ruprecht; Berlin (from D. Botting, *Humboldt and the cosmos*, Sphere Books, 1973); Gauss (from A. Von Schwiger-Lerchenfeld, *Atlas der Himmelskunde*, 1898); Queenwood (from *The Graphic*, 25th December 1880); Liouville (from *Scripta mathematica*, 1936); Collège de France (from a postcard of around 1920); Bertrand, courtesy Archives de l'Académie des Sciences de Paris; Brioschi (from *Acta mathematica*, 1912), courtesy Mittag-Leffler Institute, Stockholm; Cayley (from *Nature*, 20th September 1883) courtesy MacMillan Magazine Limited; University College (from an engraving by C. W. Radcliffe), University College School, and Hirst (seated), courtesy University College London Library, ref. College Collection; Tyndall (from *Vanity Fair*, 6th April 1872); Huxley (from *Vanity Fair*, 28th January 1871); Royal Naval College, Greenwich, The Photographic Greeting Card Co. Ltd., London; Fellows of the Royal Society (from *The Graphic*, 20th July 1889); page of Hirst's diary, courtesy Rev. Arnold Hirst; other pictures come from the collections of the Open University and the second author. While every effort has been made to secure copyright, copyright-holders who feel that their rights have been infringed should contact the second author, and a correction will appear in a later issue.

46. Proposed by H. C. WHITAKER,
A. M., Ph.D., Professor of Mathematics,
Manual Training School, Philadelphia,
Pennsylvania.

“There was an old woman tossed up in
a basket

Ninety times as high as the moon.”

Mother Goose

Neglecting the resistance of the air,
how long did it take the old lady to go
up?

American Mathematical Monthly
3, (1896) p. 281

Densest Packings of Congruent Circles in an Equilateral Triangle

Hans (J. B. M.) Melissen

1. INTRODUCTION. How large is the smallest square box that can contain n milk-bottles? If n points are distributed in a circle such that the distance between any two points is at least d , what is the largest possible value for d ? Figure 1 shows why such problems are closely related. If K is a circular disc, or a polygonal region whose edges are all tangent to a circle, packing n equal circular discs of maximum diameter inside K is equivalent to finding n points in K such that the pairwise minimum distance between points is maximal. For instance, in a unilateral triangle, these points are the centers of circles of diameter d that pack into $(1 + \sqrt{3}d)K$. We will refer to d as the *maximum separation distance* of n points in K . As there seems to be little hope of solving the packing problem for all n , research has been focussed on asymptotic estimates and on the investigation of small values of n .

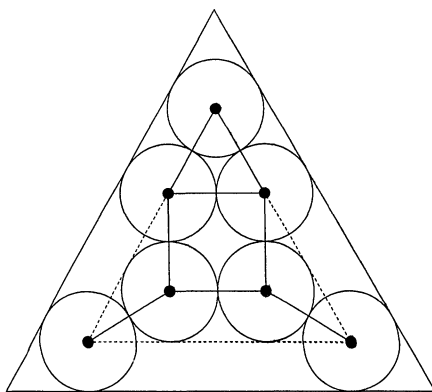


Figure 1. Densest packing of seven circles in an equilateral triangle. Seven points in an equilateral triangle with largest possible minimum distance between the points.

During the last decades much progress has been made for circle packings inside a number of simple geometrical shapes, such as the square and the circle. Solutions were found by trial and error or by computer aided optimization. Although near-optimal packings are easy to construct, few optimality proofs have appeared so far and many conjectures still rest unproven. An excellent review with relevant references can be found in [2]; see also [4, 11].

In 1969 Pirl [13] exhibited circle packings in a circle for $n = 2, \dots, 20$ and proved their optimality for $n \leq 10$. A proof for $n = 11$ was given recently by the author [8].

Optimal circle packings in a square have been constructed for $n = 6$ by Graham, for $n = 7$ by Schaer (both unpublished), for $n = 8$ by Schaer and Meir [14] and for $n = 9$ by Schaer [15]. Wengerodt (and Kirchner) [18, 17, 19, 7] gave proofs for $n = 14, 16, 25$ and $n = 36$.

Another problem that comes to mind is the packing of equal circles into an equilateral triangle. Surprisingly, only the case of the triangular numbers $n = k(k + 1)/2$ has been tackled in the literature [12, 2]. In the vein of Pirl, Schaer and Wengerodt we will provide optimal arrangements for $n \leq 10$, $n = 12$ and give an alternative proof for the triangular numbers.

The closely related problem of partitioning an equilateral triangle into subregions such that the maximum of the diameters is minimal has been studied by Graham [6]. Optimal packings of 2, 3, 4, 5, 8, 9 and 10 equal spheres in a regular tetrahedron can be found in [1].

2. OPTIMAL PACKINGS IN AN EQUILATERAL TRIANGLE. Figures 2a–k and 2p show arrangements of n points inside a unilateral triangle for which the minimum distance between the points is maximal. The solid lines in the figures connect those pairs of points for which the distance is equal to the maximum separation distance d_n . The values of d_n are given in Table 1. For $n = 2, 3, \dots, 10, 12$ we will prove that the arrangements shown are indeed optimal. The proofs for $n = 2, \dots, 7, 10$ consist in constructing a decomposition of the triangle into at most $n - 1$ subregions. Dirichlet's pigeon-hole principle tells us that one of the subregions must contain at least two points. The maximum diameter of the subregions is then an upper bound for the minimum possible distance between two points of the arrangement. In the cases under consideration, this upper bound is attained by the given configuration. The optimality proof for the arrangements of eleven points is rather involved and will be the subject of a separate paper. The cases $n = 2, 3$ are evident, so we will proceed with $n = 4$.

2.1. Arrangement of Four, Five and Six Points. Two of the four points must lie in the same subregion from the partition shown in Figure 3a, so $d_4 \leq 1/\sqrt{3}$. If the upper bound is attained, then one point must lie at the center and the other one is a vertex of the triangle. The only possible locations left for the other two points are then the other two vertices of the triangle, so for $n = 4$ the configuration is unique.

Using the partition of the triangle into four triangles as in Figure 2e, it follows that the maximum separation distance for $n = 5$ and $n = 6$ is equal to $1/2$. The configuration for $n = 5$ is just the arrangement for $n = 6$ from which one arbitrary point has been removed. This is the only freedom allowed in finding an optimal arrangement for $n = 5$.

2.2. Arrangements of Seven Points. An interesting feature of $n = 7$ is that, apart from reflected configurations, there are two different types of optimal solutions as is illustrated in Figures 1 and 2f. One is symmetric and rigid. In the other one (dashed in Figure 2f), where the interior point on the left is moved to the base of the triangle, the position of the left point in the second row is no longer unique. By projecting on the base of the triangle it can be seen from the configuration in Figure 2f that its separation distance is equal to $d_7 = (\sqrt{3} - 1)/2$. The partition shown in Figure 3b is based on the points that lie on the edges and on the bisectors of the triangle, and that are at distance d_7 from the vertices, together with the center of the triangle. From this partition it follows immediately that the maximum

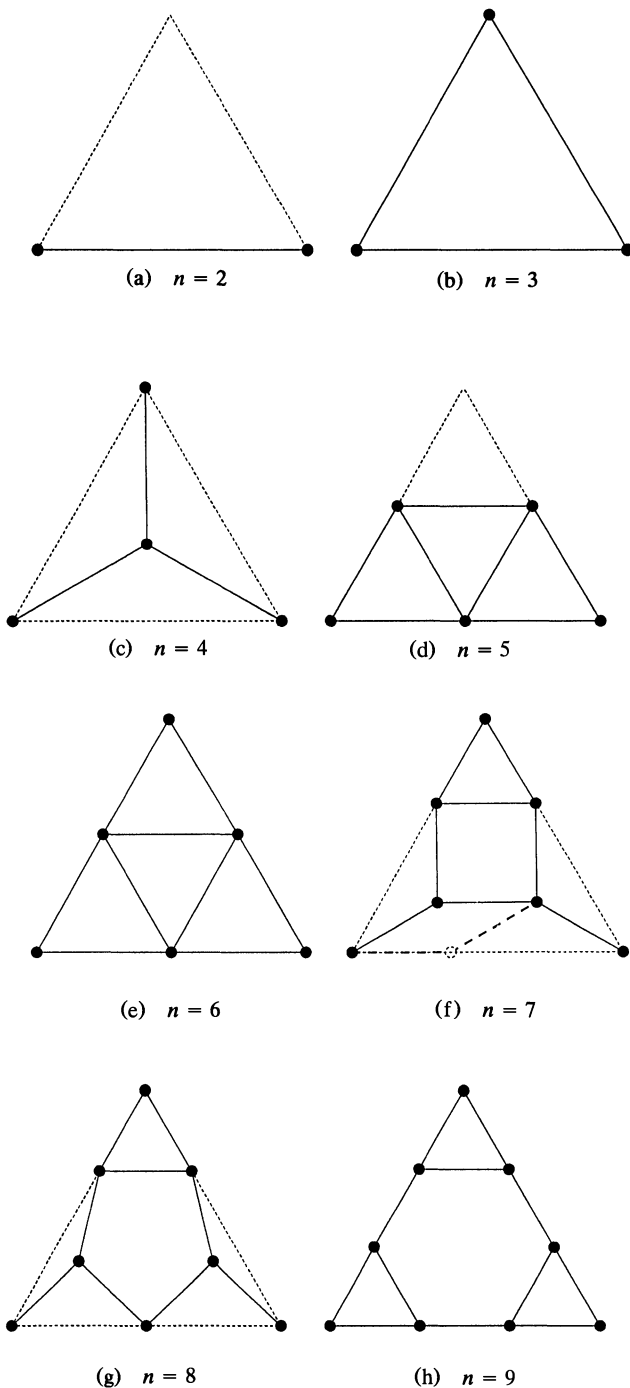


Figure 2. Optimal and conjectured optimal (*) arrangements of points in a unilateral triangle. The solid line segments are of length d_n .

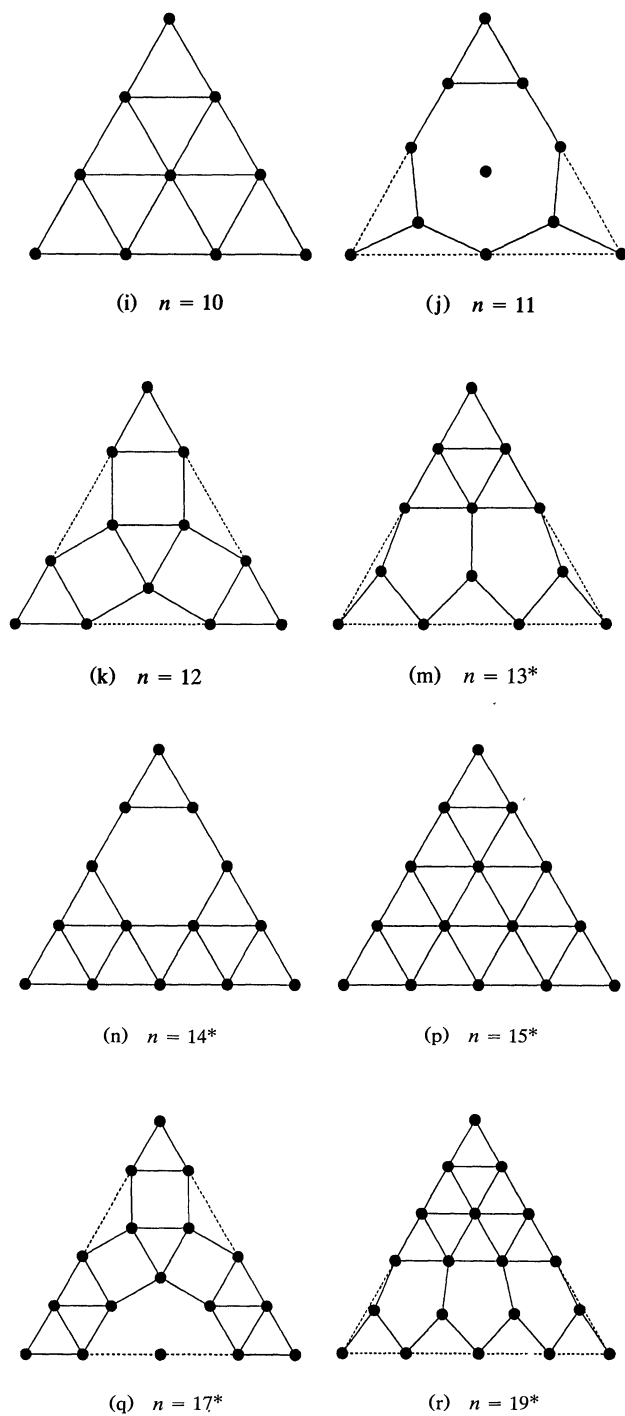


Figure 2. (Continued).

TABLE 1. Maximum separation distance d_n of n points in a unilateral triangle

n	max. separ. distance d_n		n	max. separ. distance d_n	
2, 3	1	= 1.000000 ...	12	$2 - \sqrt{3}$	= 0.267949 ...
4	$1/\sqrt{3}$	= 0.577350 ...	13*		= 0.251813 ...
5, 6	$1/2$	= 0.500000 ...	14*, 15	$1/4$	= 0.250000 ...
7	$(\sqrt{3} - 1)/2$	= 0.366025 ...	17*	$(3 - \sqrt{3})/6$	= 0.211324 ...
8	$(\sqrt{33} - 3)/8$	= 0.343070 ...	19*		= 0.200321 ...
9, 10	$1/3$	= 0.333333 ...	$k(k+1)/2 - 1^*$	$1/(k-1)$	
11	$(3 - \sqrt{6})/2$	= 0.275255 ...	$k(k+1)/2$	$1/(k-1)$	

*marks the conjectured values.

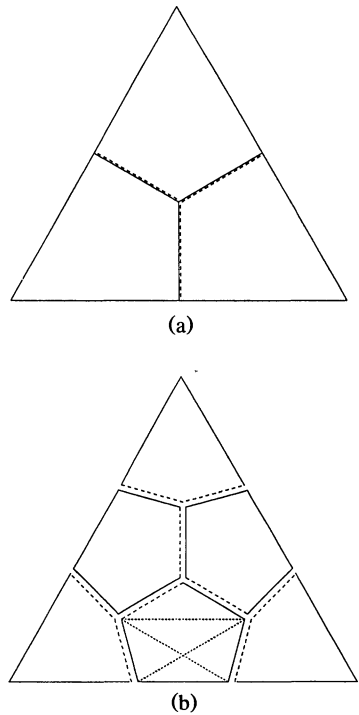


Figure 3. Partitions for $n = 4$ and $n = 7$. The solid lines indicate to which subregion each edge belongs. The three dotted lines in (b) are of length d_7 .

separation distance is equal to d_7 . The pentagonal regions can contain at most two points at distance d_7 , in exactly one way, whereas the quadrilateral regions can accommodate only one point. Easy combinatorial arguments show that only the configurations described above are possible.

2.3. Arrangements of Eight Points. A straightforward computation shows that the separation distance for the configuration in Figure 2g satisfies an equation of degree four, leading to a separation distance of $d_8 = (\sqrt{33} - 3)/8$. The arrangement is unique up to rotations.

To prove that d_8 is optimal, suppose that we have a configuration for which the distance between any two points is at least $d > d_8$. It is easy to see that a point that is closest to a vertex of the triangle can be moved to that vertex without

disturbing the optimality of the solution. We can therefore assume that the three vertices of the triangle are part of the configuration. This assumption will not restrict the total number of solutions to be found.

We will make use of the decomposition in Figure 4a. The vertices in this partition can be found by using the points from the arrangement in Figure 2g and their rotated images, together with the center of the triangle. Consider the closed region formed by the union of the subregions $R_1, R_2, R_3, Q_1, Q_2, Q_3$. In this region five points must be accommodated at a mutual distance of at least d . All its subregions have a diameter of at most d_8 . As $d_8 < d$, each cannot hold more than one point of the solution. This means that two of the Q_j , together with their interjacent R -region must each contain one point of the solution, for instance Q_1, R_1, Q_2 . This cannot happen, because $|A_3D| = |B_3D| = d_8$. Here D is the midpoint of A_1B_1 (divide $Q_1 \cup R_1 \cup Q_2$ with a cut along DC and apply the pigeon-hole principle).

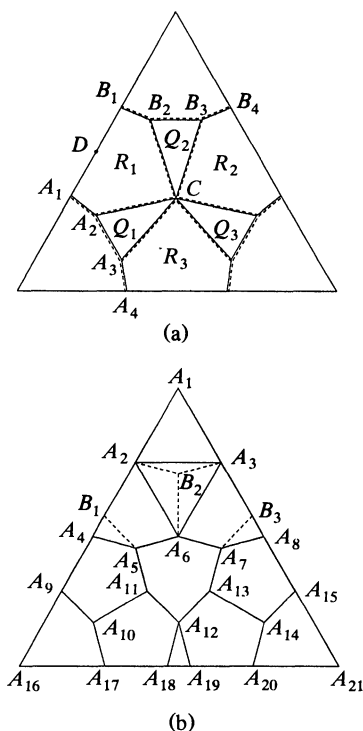


Figure 4. Partitions for $n = 8$ and $n = 12$.

Now we shall determine all possible configurations for which the separation distance is equal to d_8 . Each \bar{R}_j (the closure of R_j) can contain at most two points of the solution. For instance, for \bar{R}_1 , the possible combinations would be $A_1 - B_2$ and $A_2 - B_1$. First, we show that no Q_j can contain a point of the solution in its interior.

1. If two of the Q -regions, for instance Q_1 and Q_2 , have a point of the solution in their interior, then \bar{R}_1 cannot contain a point. Furthermore no point can be in the interior of Q_3 , otherwise there could be no solution points in \bar{R}_2 and \bar{R}_3 . This

implies that the union of \bar{R}_2 and \bar{R}_3 must contain at least three points of the solution, so one must contain two points. This is impossible, because one of these points (A_3 , A_4 , B_3 or B_4) would then be too close to at least one of the solution points in Q_1 or Q_2 .

2. If only Q_1 has a solution point in its interior, then \bar{R}_1 and \bar{R}_3 will not contain more than one point each, so there must be two points in \bar{R}_2 . This cannot occur because one of these points would be too close to the points of the configuration in \bar{R}_1 and \bar{R}_3 .

The five points must therefore be distributed over \bar{R}_1 , \bar{R}_2 , \bar{R}_3 . As the center of the triangle cannot be part of the solution, one of the three regions (e.g. \bar{R}_3) contains only one point. This implies the solution given in Figure 2g. The arrangement is unique up to rotations.

2.4. Arrangements of Nine and Ten Points. The unique configuration for $n = 10$ is an easy consequence of the pigeon-hole principle, applied to the obvious subdivision into triangles (see Figure 2i). The configurations for $n = 9$ can be obtained by removing one arbitrary point from the arrangement for $n = 10$. Unfortunately the pigeon-hole principle cannot be applied, because a partition into eight regions must contain a subregion of diameter $2/(1 + \sqrt{3} + \sqrt{6\sqrt{3}}) > 1/3$ ([6]).

First, we shall demonstrate that the maximum separation distance for $n = 9$ is equal to $1/3$. Suppose that for some configuration the distance between the points is at least $1/3 + \varepsilon$, where $\varepsilon > 0$. This means that there must be exactly one point in each subregion in Figure 2i. The three points in the three outermost triangles prohibit other points from coming within a distance ε from the edges of these triangles. Consequently, the region inside the hexagon where the remaining six points should be situated is actually contained in a disc of radius $r_0 = \sqrt{1 - 3\varepsilon + 9\varepsilon^2}/3 < 1/3$. According to Pirl [13, §2], the separation distance of these points cannot exceed r_0 , which contradicts the assumption that the distance exceeds $1/3$.

Having established d_9 it is not difficult to see that all optimal arrangements for $n = 9$ can be obtained by removing one arbitrary point from the arrangement for $n = 10$. This follows from the fact that the circumscribed circle of the six innermost triangles can enclose at most seven points with a mutual distance of at least $1/3$. On the other hand, the three regions outside this circle can contain at most three points in all, so there must be at least six points in the circular disc. From the configurations for the circle found by Pirl it follows that only the vertices of the small triangles can be part of the configuration.

2.5. Arrangements of Twelve Points. The unique optimal configuration for $n = 12$ is shown in Figure 2k. Consider the partition as indicated by the solid lines in Figure 4b. The coordinates of the nodes can be found in Table 2. The subdivision is symmetric in the bisector through A_1 . The triangle is now divided into twelve regions whose diameter is at most $d_{12} = 2 - \sqrt{3}$. If the maximum separation distance of an arrangement were larger than d_{12} , then there would be exactly one point in each subregion. The presence of a solution point in $A_{18}A_{19}A_{12}$ subsequently implies that there is a point in the interior of $A_{10}A_{11}A_{17}$, $A_4A_5A_{11}A_9$, $B_1A_5A_6A_2$ and of $A_2A_3B_2$, so there can be no point in $A_1A_2A_3$. This contradiction implies that d_{12} is the maximum separation distance.

Next, we will find the unique arrangement corresponding to this maximum separation distance. Arguments similar to those already discussed show that there can be no point in $A_{18}A_{19}A_{12}$ (with the possible exception of A_{12}). By symmetry

TABLE 2. Coordinates of the nodes in the partitions for $n = 12$

x		y	x		y
A_1	0	$\frac{1}{2}\sqrt{3}$	A_{11}	$\frac{5}{2} - \frac{3}{2}\sqrt{3}$	$-\frac{3}{2} + \sqrt{3}$
A_2	$-1 + \frac{1}{2}\sqrt{3}$	$\frac{3}{2} - \frac{1}{2}\sqrt{3}$	A_{12}	0	$1 - \frac{1}{2}\sqrt{3}$
A_4	$-2 + \sqrt{3}$	$3 - \frac{3}{2}\sqrt{3}$	A_{16}	$-\frac{1}{2}$	0
A_5	$-1 + \frac{1}{2}\sqrt{3}$	$-\frac{1}{2} + \frac{1}{2}\sqrt{3}$	A_{17}	$\frac{3}{2} - \sqrt{3}$	0
A_6	0	$3 - \frac{3}{2}\sqrt{3}$	A_{18}	$-\frac{7}{2} + 2\sqrt{3}$	0
A_9	$\frac{1}{2} - \frac{1}{2}\sqrt{3}$	$-\frac{3}{2} + \sqrt{3}$	B_1	$\frac{3}{2} - \sqrt{3}$	$-3 + 2\sqrt{3}$
A_{10}	$-2 + \sqrt{3}$	$1 - \frac{1}{2}\sqrt{3}$	B_2	0	$-2 + \frac{3}{2}\sqrt{3}$

the same must be true for $A_4A_5B_1$ and $A_7A_8B_3$. Now we adapt the decomposition of the triangle to one with the three bisectors of the triangle as axes of symmetry. This is indicated by the dashed line segments in Figure 4b. The region $A_4A_5A_{11}A_{12}A_{18}A_{16}$ with the segments A_5A_{11} and $A_{11}A_{12}$ excluded can contain a maximum of four points of the optimal arrangement, and this in exactly one way ($A_4, A_{10}, A_{16}, A_{18}$). This is evident from the partition into three subregions of diameter d_{12} . The central hexagon can contain a maximum of three points (A_5, A_7, A_{12}). Straightforward combinatorial arguments then show that three solution points in the hexagonal region correspond to the arrangement in Figure 2k, whereas two or less points cannot lead to a solution.

2.6. Arrangements for Triangular Numbers. For the triangular numbers $n = k(k+1)/2$, ($k \geq 2$), the obvious candidates for the optimal arrangements are given by the regular triangular lattice arrangement in analogy to Figures 2b, e, i, p. Unfortunately, the partitioning trick is unsuitable to prove this for all triangular numbers. This is because the number of triangles $(k-1)^2$ exceeds the number of points n , for $k \geq 5$. Oler [12] asserted that the minimum distance between $n+1$ points in a unilateral triangle is smaller than $1/(k-1)$. Looking at his proof we notice that Oler actually proved that $d_n = 1/(k-1)$, however, without showing that the obvious arrangement is indeed unique. The proof is based on a general inequality that was conjectured by Zassenhaus and proved by Oler in 1961 (see [5]). This inequality provides an upper bound for the number of points n that can be placed in a planar convex compact set K at a mutual distance of at least 1, expressed in terms of the area $\mu(K)$ and the perimeter $\mu(\partial K)$ of K :

$$n \leq \frac{2}{\sqrt{3}}\mu(K) + \frac{1}{2}\mu(\partial K) + 1. \quad (1)$$

The optimality proof for the triangular numbers is obtained by applying this inequality to an equilateral triangle. A similar inequality of Groemer (1960, see [10]) can also be used. We shall give a more straightforward proof by deriving this inequality directly for the case of a unilateral triangle. In addition we can also conclude the uniqueness of the optimal solution.

Theorem. *If $n \geq 2$ points are placed inside a unilateral triangle then the minimum of the mutual distances between these points, d , satisfies the following inequality:*

$$d \leq \frac{2}{\sqrt{8n+1}-3}. \quad (2)$$

Equality is attained only if $n = k(k+1)/2$, ($k \geq 2$), for points on a regular triangular lattice.

Proof: Suppose that for some n an arrangement is given. The circles centered around these points with radius $r = d/2$ then form a packing inside an equilateral triangle with a side-length of $1 + 2\sqrt{3}r$. The plane could be tiled with these triangles to obtain a global circle packing. For our purpose, however, this packing is not good enough. We will use a more economical packing shown in Figures 5 and 6. The packing is reflected in a line parallel to one side of the triangle

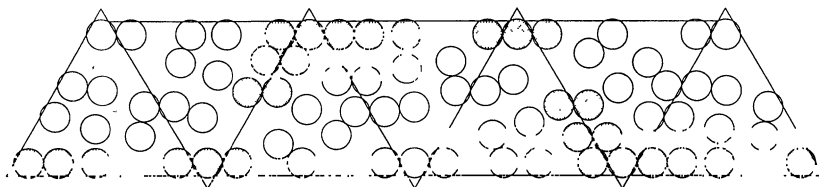


Figure 5. Packing of packed triangles in a strip.

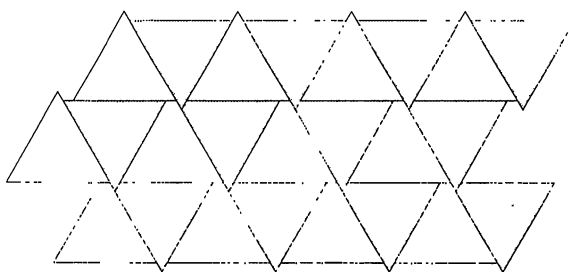


Figure 6. Tiling with truncated triangles.

touching the circles. This mirror image is then slightly moved until it fits snugly into the original arrangement (see Figure 5). It is easy to see that this can always be done. This process is repeated to obtain an infinitely long strip of circle packings. The same technique is then applied to the strip resulting in a global circle packing. This is possible because in the strip the arrangement of circles repeats itself after six triangles. The side-length of the triangles in Figure 5 is $1 + 3r$. Although these triangles overlap, the plane can be tiled by the trapezoids as shown in Figure 6 (the shaded regions correspond to mirror images of the arrangement). It is a well-known result of Thue [16, 4] that the density of a plane circle packing cannot exceed $\pi/\sqrt{12}$ and that this maximum value is attained for the honeycomb packing where each circle touches six neighbors and the centers are on a regular triangular lattice. This implies the following estimate:

$$\frac{n\pi r^2}{\frac{1}{4}\sqrt{3}[(1+3r)^2 - r^2]} \leq \frac{\pi}{2\sqrt{3}},$$

which leads to inequality (2). For triangular numbers $n = k(k+1)/2$, this inequality reduces to $d \leq 1/(k-1)$; in this case the hexagonal packing is the unique optimal solution. ■

3. CONJECTURES. For $n = 13, 14, 19$, conjectures for the optimal arrangements are presented in Figure 2m, n, r. The optimal arrangements for $n = k(k+1)/2 - 1$ seem to be obtained by removing one arbitrary point from the arrangement for $n = k(k+1)/2$. This conjecture was posed as an open problem by Erdős and Oler

[12, 2]. We have already shown its validity for $n = 2, 5, 9$. The conjecture actually implies a still open conjecture of Fejes Tóth [3], which states that if $n + 1$ circles are removed from the honeycomb packing of equal circles, and n are packed again in the resulting interstitial space, then we always end up with the original packing from which one circle has been removed.

The configurations in Figures 2c, g, m, r suggest a possible form for the optimal arrangements for $n = k(k + 1)/2 - 2$, ($k \geq 3$). First $k - 3$ layers of $(k - 3)^2$ equilateral triangles, followed by a layer of $k - 3$ pentagons. We conjecture that these are the unique optimal configurations in these cases (up to rotations). The conjecture is true for $n = 4$ and 8.

Conjectures for $n = 16, 17$ and 18 are presented in [9]. One configuration is shown in Figure 2q.

Acknowledgments. I would like to thank Prof. J. H. van Lint, Dr. P. C. Schuur, Dr. A. J. E. M. Janssen and Drs. M. J. J. B. Maes for reading the drafts of the paper and for useful discussions.

REFERENCES

1. K. Bezdek, Densest packing of a small number of congruent spheres in polyhedra, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* **30** (1987) 177–194.
2. H. T. Croft, K. J. Falconer and R. K. Guy, *Unsolved Problems in Geometry*, Springer Verlag, Berlin, 1991, 107–111.
3. L. Fejes Tóth, Solid circle-packings and circle-coverings, *Studia Sci. Math. Hung.* **3** (1968) 401–409.
4. ———, *Lagerungen in der Ebene, auf der Kugel und im Raum*, (1953) 2^e Auflage (1972), Springer Verlag, Berlin.
5. J. H. Folkman and R. L. Graham, A packing inequality for compact convex subsets of the plane, *Canad. Math. Bull.* **12** (1969) 745–752.
6. R. L. Graham, On partitions of an equilateral triangle, *Canad. J. Math.* **19** (1967) 394–409.
7. K. Kirchner and G. Wengerodt, Die dichteste Packung von 36 Kreisen in einem Quadrat, *Beiträge Algebra Geom.* **25** (1987) 147–159.
8. J. B. M. Melissen, Densest packings of eleven congruent circles in a circle, to appear, *Geom. Dedicata*.
9. J. B. M. Melissen and P. C. Schuur, Packing 16, 17 and 18 circles in an equilateral triangle, to appear, *Discrete Math.*
10. J. Molnár, On the packing of unit circles in a convex domain, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* **22/23** (1979/80) 113–123.
11. W. O. Moser and J. Pach, *Research Problems in Discrete Geometry*, Montreal (1984).
12. N. Oler, A finite packing problem, *Canad. Math. Bull.* **4** (1961) 153–155.
13. U. Pirl, Der Mindestabstand von n in der Einheitskreisscheibe gelegenen Punkten, *Math. Nachr.* **40** (1969) 111–124.
14. J. Schaer and A. Meir, On a geometric extremum problem, *Canad. Math. Bull.* **8** (1965) 21–27.
15. J. Schaer, The densest packing of 9 circles in a square, *Canad. Math. Bull.* **8** (1965) 273–277.
16. A. Thue, Om nogle geometrisk-taltheoretiske theoremer, *Forhdl. Skand. Naturforskere* **14** (1892) 352–355.
17. G. Wengerodt, Die dichteste Packung von 16 Kreisen in einem Quadrat, *Beiträge Algebra Geom.* **16** (1983) 173–190.
18. ———, Die dichteste Packung von 14 Kreisen in einem Quadrat, *Beiträge Algebra Geom.* **25** (1987) 25–46.
19. ———, Die dichteste Packung von 25 Kreisen in einem Quadrat, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* **30** (1987) 3–15.

Philips Research Laboratories
P.O. Box 80.000
5600 JA Eindhoven
The Netherlands
melissen@prl.philips.nl

Partnerships¹

Alan H. Schoenfeld

In the section called “Action,” *Everybody Counts* (National Research Council, 1989) issued this clarion call:

“In the next decade, the United States has a historic opportunity to revitalize mathematics education . . .

“There are at this time both a particular urgency and a special opportunity for reform of mathematics education. Since mathematics is the foundation of science and technology, reform is needed to prepare the more highly skilled work force that the nation now needs. Because of the emerging general agreement within the mathematics, mathematics education, and related professional communities on goals for mathematics education and means for achieving them, there is at this time a special opportunity for the nation to push ahead boldly in this area of education. (page 87)”

The mathematics education community has indeed been pushing boldly ahead, and it is of great interest to note the character of the advances—especially as they contrast with the character of the field in its early days, approximately a quarter-century ago.² For example, Joe Crosswhite recalls that the first research sessions at an annual NCTM meeting were held behind the stage, behind a closed curtain—placed by conference organizers at a safe physical and psychological distance from more “teacherly” conference activities. Physically and intellectually, the research community stood apart. Indeed, its apartness was manifested in multiple ways: in focus, in methods, and in the communities from which it drew. As in all of the social sciences through the 1960’s and 1970’s, the methods employed tended to be “rigorous” and “scientific,” with a focus on experimental studies and statistical analyses. Many experiments took place in the lab, at some remove from instruction. Those studies which took place in classrooms tended to downplay the complexity of classroom interactions, focusing on specific instructional “variables” and their effects, as determined statistically. Hence in 1978 Kilpatrick felt obliged to suggest that educational researchers might have lost sight of meaningful mathematical behavior in their search for methodological rigor, and that the community might have much to learn from unrigorous but interesting studies such as the

¹This report was prepared by Alan H. Schoenfeld, University of California at Berkeley, chair of the NCTM Research Advisory Committee, and was reviewed by members of the Committee. At the time this report was prepared in April 1993, committee members were Deborah Ball, Michigan State University; Robert Davis, Rutgers University; Beverly Ferrucci, Keene State College of New Hampshire; Marilyn Hala (Staff Liaison), NCTM Headquarters; Miriam Leiva (Board Liaison), University of North Carolina at Charlotte; Susan Jo Russell, TERC; William Tate, University of Wisconsin. Reprinted with permission from the JRME, copyright 1993, by the National Council of Teachers of Mathematics.

²Papers in mathematics education can be traced back a good many years, of course, but the creation of the *Journal for Research in Mathematics Education* about 25 years ago is generally taken as a sign of the coalescence of the discipline.

largely qualitative teaching experiments carried out in the Soviet Union by researchers such as Krutetskii (1976). In terms of communication across communities, Pólya was the exception that *probed* the rule: after the burst of energy that produced the New Math, there was little interaction between the mathematics and the math-ed communities, especially along the lines of research.

Things have changed! As noted in *Everybody Counts*, “real change requires action by everyone involved in mathematics education” (page 93). The Mathematical Sciences Education Board, formed in 1985, represents an attempt to bring together the various constituencies that have a stake in mathematics education. Multiple communities have a stake in getting things right. More importantly, multiple communities have major contributions to make.

Mathematicians, for example, live and breathe the discipline; they can offer a deep sense of what it is to engage in mathematics, and a sense of what might be called the “mathematical validity” of a curriculum—whether the ideas and processes with which students engage tend to reflect the deep underlying notions of mathematical “doing.” In recent years the mathematical community’s interest in educational issues has mushroomed: witness the existence of Mathematicians and Educational Reform, a grass roots organization of university mathematicians with interest in contributing to K-12 mathematics education, and the fact that the American Mathematical Society has created a Committee on Education, one major function of which is to establish liaison with other, longer-established groups with educational interests.

In many ways the teaching community has been galvanized by the *Curriculum and Evaluations Standards for School Mathematics* (NCTM, 1989) and the *Professional Standards for Teaching Mathematics* (NCTM, 1991). The wisdom of the profession was a major factor in the creation of those documents, and will be an essential resource if we are to reach to goals set forth in them: teachers live the reality of instruction in their classrooms, and must be the wellspring of the reform movement. And the professional teaching community is ready for interactions with the other communities, as evidenced by the spectacular growth in NCTM membership and attendance at annual NCTM meetings in recent years, and the diversification of conference programs to include a significant focus on research-related activities.

Beyond the classroom, schools, school districts, parental understanding and influence, state departments of education and national curricular influences (texts and tests) are major factors that affect the ways in which reform can take place, and whether it will be sustained. Members of all these communities need to be enfranchised, and need to contribute to dialogue and change.

Last but not least, the community of mathematics educators has grown spectacularly over the past 25 years, and is capable of being a central “team player” in the reform of the profession. Even a cursory glance at the *Handbook of Research on Mathematics Teaching and Learning* (Grouws, 1992) reveals how vibrant and robust an enterprise research in mathematics education has become. A closer look reveals how much the field has broadened, in the range of methods it employs and the phenomena it explores. Methods include computer simulations of individual cognition, clinical interviews, classic laboratory studies, ethnographic analyses of classroom cultures, qualitative studies of teacher and student beliefs and their effects on behaviors, and more. The classroom, once seen by most as “too complex” for careful studies of mathematical thinking and learning, is now seen by many as the natural place for such studies. Along with inward growth came an outward look: the mathematics education community now looks to teachers, mathematicians,

psychologists, cognitive scientists, anthropologists, and numerous other communities for issues, ideas, and inspiration as it seeks to grapple with the complex phenomena of mathematical understanding, thinking, and learning.

We are, then, at an important point in the development of mathematics education. There is general recognition that the problems we face are large, and that they require the concerted effort of all the major constituencies involved in the educational process. Although many of those constituencies have in the past been communities apart, there is now unprecedented potential for collaborative work and joint community building. Over the past few years, the Research Advisory Committee in particular and NCTM in general have been moving in those directions. Here are some examples of recent, proposed, and potential projects.

Two years ago (July 1991) RAC reported on the NCTM *Standards* Research Catalyst conferences, which were then in progress. One major goal of the conferences, supported by the NSF and held in March and December 1991, was to focus research on major themes in the *Curriculum and Evaluations Standards for School Mathematics*. The profession needed to know more about assessment, curriculum change, communication, policy, representational tools and models, and the changing secondary curriculum; it made sense to have focus groups address those issues. But an equally important goal was the enfranchisement of a new research community, reaching out from the traditional base of mathematics educators to teachers, administrators, and others to begin research and research partnerships in these areas. By any measure, the effort was a significant success: a number of new researchers received NSF seed grants for work stimulated by the conference, and some of the partnerships formed (e.g. the communications group) continue today as active research collaboratives.

We hope, a few years from now, to report on a similar undertaking related to the *Professional Standards for Teaching Mathematics* entitled the "Collaborative for enhancing research in mathematics teaching." The goal of the proposed collaborative is to build a community of people working together to conceptualize and carry out research on mathematics teaching, with an emphasis on broadening the kinds of research being used to inform reform efforts and exploring new ways to communicate about research to diverse audiences. The collaborative is especially interested in attracting new researchers, experienced researchers new to research on mathematics teaching, mathematicians, and mathematics educators whose activities have not traditionally been considered to be research (e.g. classroom teachers, staff developers, administrators, college teachers).

With the help of the Exxon Education Foundation, work is now under way on the first phases of a project entitled "Recognizing and recording reform in mathematics education: Documenting the effects of the National Council of Teachers of Mathematics *Curriculum and Evaluations Standards* and *Professional Standards for Teaching Mathematics*." This project, quite large in scope, is intended to take a systemic view of change, and to help the community at large understand the dynamics of educational reform. This project will, of necessity, involve all the major constituencies involved in mathematics education. From the project description: "Such a project, through its structure and intent, emphasizes that the changes outlined in the *Standards* documents will not happen quickly, or easily, or without experimentation and false starts. A project such as this confirms that it is not only acceptable, but essential, to learn from the process of implementation and change and to disseminate and share that knowledge openly, even

though the stories that emerge will describe obstacles and difficulties as well as successes.”

Finally, a set of activities on “Partnerships in research” is in the planning stages. The task force working on the project expects to assemble videotapes of classroom instruction that can serve as the focal points for conversations among mathematicians, teachers, administrators, and mathematics education researchers regarding the values, goals, and practices of mathematics instruction. It hopes that first at a national conference, and then at a series of spin-off local conferences, the videotapes and related support materials will serve as means of facilitating conversations among those groups, all of which are essential for continued progress in educational reform.

These are exciting times. The spirit of reform is in the air; the communities necessary to promote it are open to collaboration; and efforts to join forces in this important collaborative enterprise are being undertaken. That the various communities listed above have grown to the point where they recognize their interdependencies and are willing to build partnerships bodes well for all concerned, and should cheer us all—but it should not leave us feeling complacent. We have just embarked on the collaborative trail, and there is much more to be done. Although one can point to exceptions in individual states and locales, the research community has not, in general, been adequately engaged with policy makers at the state and national levels. Local, state, and national policies may or may not be consistent with our best understandings. Likewise, local, state, and national assessment measures may support or may undermine what we would like to have happen in our nation’s mathematics classrooms. Much more direct contact and productive interaction among the policy, assessment, and research communities is necessary. Similarly, although there are encouraging signs of interactions, the research community has yet to engage adequately with issues of teacher preparation. And, of course, this brief list of necessary collaborations can be expanded without difficulty. In sum, let us take pleasure in the progress we have made. Then, let us return to the task of making and strengthening essential partnerships for progress in mathematics education.

REFERENCES

- Grouws, D. (Ed.) (1992). *Handbook of research on mathematics teaching and learning*. New York: MacMillan.
- Kilpatrick, J. (1978). Variables and methodologies in research on problem solving. In L. Hatfield (Ed.), *Mathematical problem solving* (pp. 7–20). Columbus, OH: ERIC.
- Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children*. (J. Teller, translator; J. Kilpatrick, & I. Wirszup, Eds.). Chicago: University of Chicago Press.
- National Research Council. (1989). *Everybody counts: A report to the nation of the future of mathematics education*. Washington, DC: National Academy Press.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. (Reston, VA: NCTM.
- National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.
- Research Advisory Committee (1991). NCTM Standards research catalyst conference. *Journal for Research in Mathematics Education*, 22(4), 293–296.

Education and Mathematics Department
University of California, Berkeley
Berkeley, CA 94720

A Simple Proof of Pascal's Hexagon Theorem

Jan van Yzeren

Pascal's Theorem. *If the vertices of a hexagon lie on a circle and the three pairs of opposite sides intersect, then the three points of intersection are collinear.*

This theorem was published in 1640 by sixteen-year-old Blaise Pascal. His original proof has been lost, and at times one wonders whether one or another of the known proofs is, in fact, Pascal's original one. This also applies to the simple proof given here.

Begin with the hexagon A_i , $i = 0, \dots, 5$ of Figure 1, and consider the circle through the points A_1 , A_4 and P_1 , where the first two points are (opposite) vertices, and the last is one of the "Pascal points" connected to them. This circle meets A_0A_1 and A_3A_4 at B_0 and B_1 respectively, and one uses arcs of the circles shown to find equal angles inscribed in them (or supplementary angles inscribed in opposite arcs). As a consequence, the triangles $P_1B_0B_1$ and $P_2A_0A_3$ have respectively parallel sides, that is, they are perspective from the point P_0 . Therefore, P_0 , P_1 and P_2 are collinear.

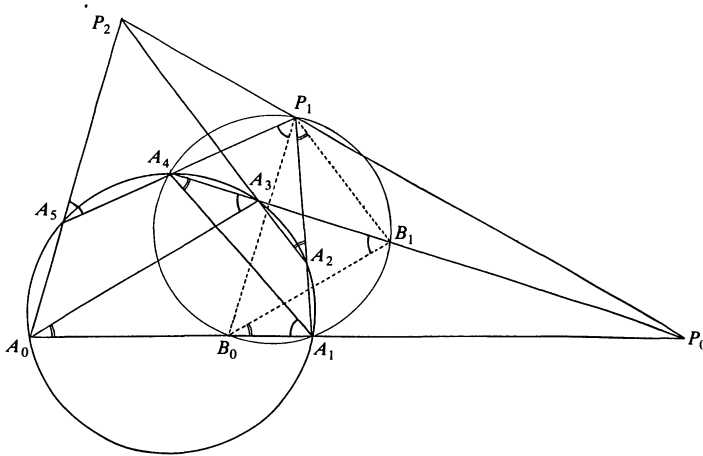


Figure 1

The proof also covers the case of $A_0A_1 \parallel A_3A_4$ (i.e., P_0 at infinity). Then, the triangles are translative, that is, P_1P_2 is parallel with A_0A_1 and A_3A_4 . The only special case not covered by the proof concerns hexagons inscribed in a circle with parallels as opposite sides. This case, however, follows easily from appropriate arcs.

Whether Pascal gave this proof is open to debate, but it seems that this proof has not turned up for 350 years. On this point Professor Coxeter kindly has commented as follows: "It is indeed remarkable that this elegant proof was not

found in 350 years, and also somewhat remarkable that Guggenheimer came close to it in 1967 and then felt obliged to introduce a peculiar lemma.” [3]

Anyway, the historic delay justifies some special attention for the heuristics of this simple proof.

The basic figure consists of two pencils of four lines joining points on a circle, viz. (Figure 2) A_0 and A_4 with, respectively, A_1, A_2, A_3 and A_5 .

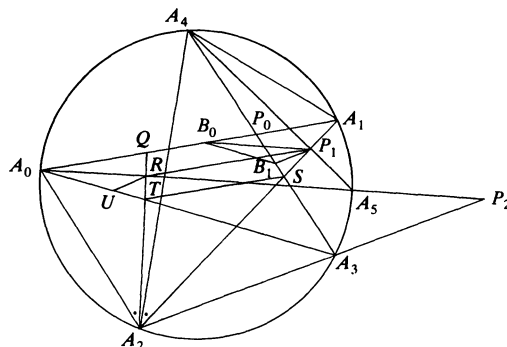


Figure 2

Evidently, the two pencils are congruent (equal angles between corresponding lines). Therefore, if $\triangle A_0A_2Q$ is made similar to $\triangle A_4A_2A_1$, the segments A_2Q and A_2A_1 are divided proportionally and $A_1A_0 \parallel P_1R \parallel ST$. Now, the crucial idea is to build up this basic figure in a converse manner, starting with two given similar triangles: $\triangle A_0A_2Q \sim \triangle A_4A_2A_1$ and forgetting the circle.

Then, choose P_1 and R on, respectively, A_1A_2 and QA_2 , such that $P_1R \parallel A_1A_0$. Similarly S and T . Hereupon the following points are defined: $A_5 = A_4P_1 \cap A_0R$, $A_3 = A_4S \cap A_0T$, $P_0 = A_4S \cap A_0A_1$, $P_2 = A_0R \cap A_2A_3$.

To prove that P_0, P_1 and P_2 are collinear:

Consider $\triangle RA_0U$, $RU \parallel P_2A_3$, and its translative image $\triangle P_1B_0B_1$. Then, B_0 lies on P_0A_0 as $P_0A_0 \parallel P_1R$, and B_1 lies on P_0A_3 , because $P_1B_1 = RU = A_2A_3 \cdot RT/A_2T = A_2A_3 \cdot P_1S/A_2S$. Therefore, the triangles $P_1B_0B_1$ and $P_2A_0A_3$ are perspective from the point P_0 and, indeed, P_0, P_1 and P_2 are collinear.

Afterwards the crucial points B_0 and B_1 can be found directly. In fact, they lie on the circumcircle of $\triangle P_1A_1A_4$, because $\angle P_1B_0A_1 = \angle A_5A_0A_1 = \angle A_5A_4A_1 = \angle P_1A_4A_1$ and $\angle A_4B_1B_0 = \angle A_4A_3A_0 = \angle A_4A_1A_0 = \angle A_4A_1B_0$. Actually, drawing the circumcircle of $\triangle P_1A_1A_4$ is the very point of the new proof.

Background of the heuristics is the fact that the metric of the Euclidean plane can be defined by giving a pair of similar triangles. After that, all other metric properties must follow by means of parallels and proportionalities (affine tools).

REFERENCES

1. H. S. M. Coxeter and S. L. Greitzer, *Geometry Revisited*, New Mathematical Library, Random House and The L. W. Singer Company, 1967.
2. H. S. M. Coxeter, *The Real Projective Plane* (2nd ed.), Cambridge University Press, 1961.
3. H. W. Guggenheimer, *Plane Geometry and Its Groups*, Holden Day Inc., Cambridge, 1967.
4. R. A. Johnson, *Advanced Euclidean Geometry*, Dover Publications, New York, 1960.
5. N. A. Court, *College Geometry*, Barnes & Noble, New York, 1959.

Department of Mathematics
Technological University Eindhoven
The Netherlands, 5600 MB

NOTES

Edited by: John Duncan

The Mathematical Relationship Between Kepler's Laws and Newton's Laws

Andrew T. Hyman

1. INTRODUCTION. Whenever a new scientific theory comes down the pike, it is greeted by skeptics who demand proof that the new theory is as good as the theory it would displace. That is why “the major scientific problem of the [seventeenth] century” was to prove that Isaac Newton’s law of gravity gives the same correct results as the older laws of Johannes Kepler [4]. This famous mathematical problem is solved below in an innovative way that requires no trigonometry, only elementary calculus, and none of the usual “clever tricks” [8].

Supposing that planets move according to Kepler’s Laws (which are reviewed in Section 2 below), then it follows that planetary acceleration is given by Newton’s central inverse-square equation (which is equation twelve below). This historic theorem was first proved by Newton, who thereby established his law of gravity as a respectable successor to Kepler’s Laws. This same theorem is proved in Section 3, using simple and straightforward methods. The reverse theorem, according to which the central $1/R^2$ equation requires Keplerian orbits, is proved in Section 4.

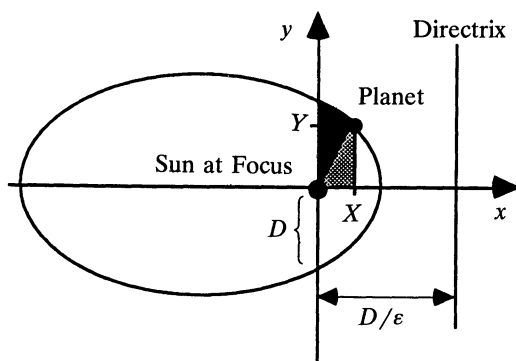
The two theorems proved here were first published in Newton’s 1687 *Philosophiae Naturalis Principia Mathematica*, or *Principia* for short. Newton admitted that the *Principia* is purposely “abstruse” ([3], p. 90), and a controversy persists as to whether Newton’s proofs are entirely legitimate ([2], p. 30). Unlike the *Principia*, the brief proofs below are quite transparent.

Kepler’s Laws are differentiated in Section 3 using only Cartesian coordinates, and this novel Cartesian approach contrasts with the usual technique of transforming to polar coordinates. Although the converse proof of Section 4 is fundamentally the same as those of a few other authors ([5], p. 178 of [11], and p. 625 of [1]), each step in Section 4 follows naturally and inexorably from what precedes it. No rabbits are pulled out of hats. The method of Section 4 is thus presented in a clear manner which compares favorably to the more common methods of solving the same problem, and also to various uncommon methods which are discussed in [10].

2. REVIEW OF KEPLER’S LAWS. Kepler deduced his laws from data supplied by the astronomer Tycho Brahe. Kepler’s Laws are:

- I. *Each planet moves along an ellipse with the Sun at a focus.*
- II. *The line from a planet to the Sun sweeps out equal areas in equal times.*
- III. *The square of a revolution’s duration, divided by the cube of the orbit’s greatest width, is the same for all planets.*

Ellipses are, of course, the closed curves formed by intersecting a cone and a plane. They were studied by the ancient Greeks (see p. 119 of [6]) who proved that the distance to a point (the “focus”) divided by the distance to a line (the “directrix”) is a constant “eccentricity” ε . A beautiful proof of this focus-directrix property was devised in 1822 by G. P. Dandelin. Dandelin’s proof appears at p. 546 of [9], and it applies to both closed ($0 \leq \varepsilon < 1$) and open ($\varepsilon \geq 1$) conic sections.



Kepler’s Laws can be translated into equations by picturing a planet as a point-particle in the x - y plane, having coordinates (X, Y) at time t (see Figure). The Sun is located at the origin, and the planet’s directrix is perpendicular to the x -axis at a distance D/ϵ from the Sun. “ D ” is called the “semi-latus-rectum” of the conic section. According to Kepler’s First Law, the distance $R \equiv \sqrt{X^2 + Y^2}$ from the planet to the Sun is given by:

Kepler's Second Law can be formulated in similarly simple terms. If the planet crosses the y -axis at time t_0 , then the area swept between t_0 and t equals the area under the curve minus the triangular area beneath the line from Sun to planet. Hence, at all times,

where “ C ” is the constant ratio of area swept to time elapsed (a new constant t_0 must be introduced whenever the planet crosses the x -axis).

$$C^2/D = K \quad (3)$$

where the constant “ K ” is the same for all planets. In summary, Kepler’s Laws are (1), (2), and (3).

3. PROOF OF CENTRAL $1/R^2$ EQUATION. Kepler’s Laws will now be used to find the acceleration of a planet. Differentiating (1) produces:

$$\frac{1}{R} \left[X \frac{dX}{dt} + Y \frac{dY}{dt} \right] = -\varepsilon \frac{dX}{dt}. \quad (4)$$

Differentiating (2), using the Fundamental Theorem of Calculus, gives:

$$Y \frac{dX}{dt} - X \frac{dY}{dt} = 2C. \quad (5)$$

A bit of algebra applied to (4), (5), and (1) makes it clear that the two velocity components are:

$$\frac{dX}{dt} = \frac{2C}{D} \cdot \frac{Y}{R} \quad (6)$$

and

$$\frac{dY}{dt} = -\frac{2C}{D} \cdot \frac{X}{R} - \frac{2C\varepsilon}{D}. \quad (7)$$

Differentiating (5) yields:

$$Y \frac{d^2X}{dt^2} - X \frac{d^2Y}{dt^2} = 0. \quad (8)$$

Differentiation of the right-hand-side of (6) is facilitated by the following identity:

$$\frac{d}{dt} \left[\frac{Y}{R} \right] = \frac{X}{R^3} \cdot \left[X \frac{dY}{dt} - Y \frac{dX}{dt} \right]. \quad (9)$$

This identity is based solely upon the definition of R .[†] By differentiating (6) and plugging in (9), (5) and (3) one gets:

$$\frac{d^2X}{dt^2} = \frac{-4KX}{R^3}. \quad (10)$$

By (8) and (10),

$$\frac{d^2Y}{dt^2} = \frac{-4KY}{R^3}. \quad (11)$$

Equations (10) and (11) can be written compactly in terms of vectors.

$$\frac{d^2\vec{R}}{dt^2} = \frac{-4K\vec{R}}{R^3}. \quad (12)$$

Equation (12) is Newton’s central inverse-square equation. This equation expresses Newton’s law of gravity for the special case where planetary mass is negligible.

[†]Incidentally, note that $[X\dot{Y} - Y\dot{X}]$ is twice the areal speed (i.e., $R^2\dot{\theta}$ in polar coordinates), where dots denote differentiation. The referee has keenly observed that therefore equation (9) is basically $[\sin \theta]' = [\cos \theta]\dot{\theta}$.

4. RECOVERY OF KEPLER'S LAWS. It remains to be seen whether a bounded orbit could satisfy (12) if it is not Keplerian. In other words, could a planet be accelerating according to (12), and yet violate Kepler's Law? It will now be proved that such an orbit is impossible, by recovering Kepler's Laws from (12). By the way, it is taken for granted that motion is confined to a plane, though this assumption is easily justified ([7], p. 105).

Equations (10) and (11) lead to (8), and integrating (8) retrieves (5) and (2). Plugging (5) into the crucial identity (9) gives:

$$\frac{d}{dt} \left[\frac{Y}{R} \right] = \frac{-2CX}{R^3}. \quad (13)$$

On account of (13) and (10),

$$\frac{Y}{R} = \frac{C}{2K} \cdot \frac{dX}{dt} + A \quad (14)$$

where "A" is a constant of integration.

The identity (9) has been very useful here, and it would have been necessary to pull this identity out of thin air were it not for the context provided by Section 3. In this context, the identity (9) has arisen in a natural way (whereas other authors have indeed pulled this identity from out of the blue).

Interchanging "X" and "Y" in (9) produces another identity which together with (5) yields:

$$\frac{d}{dt} \left[\frac{X}{R} \right] = \frac{2CY}{R^3}. \quad (15)$$

So, by (11),

$$\frac{X}{R} = \frac{-C}{2K} \cdot \frac{dY}{dt} + B \quad (16)$$

where "B" is another constant. Plugging (14) and (16) into (5) yields:

$$\frac{C^2}{K} + AY + BX = R. \quad (17)$$

If $A = B = 0$, this describes a circle. If not, (17) represents a conic section with focus at the origin, eccentricity $[A^2 + B^2]^{1/2}$, and directrix given by:

$$\frac{C^2}{K} + Ay + Bx = 0. \quad (18)$$

This interpretation of (17) follows from a simple fact of analytic geometry: the distance from a point (x_0, y_0) to a line $ax + by + c = 0$ is equal to $|ax_0 + by_0 + c|/[a^2 + b^2]^{1/2}$. This well-known fact can also be applied to (18) in order to find the distance from focus to directrix, and it is thus evident that the focus-directrix distance is as described by (3). Consequently, if Newton's central inverse-square equation holds true then all bounded orbits must satisfy Kepler's Laws, which was to be demonstrated.

ACKNOWLEDGMENTS. I thank Dr. David Griffiths of Reed College for his help. I am also grateful to the referee for suggesting a number of improvements. Furthermore, I would like to express appreciation to the European Journal of Physics for printing an article similar to this one in July of 1993 (vol. 14, no. 4).

REFERENCES

1. R. Abraham and J. Marsden, *Foundations of Mechanics*, Benjamin, Reading, Massachusetts, Second Edition, 1978.
2. V. I. Arnol'd, *Huygens & Barrow, Newton & Hooke*, Birkhäuser, Basel, 1990.
3. G. Christianson, *In the Presence of the Creator; Isaac Newton and His Times*, Free Press, New York, 1984.
4. I. Cohen, essay in *Physics* by P. Tipler, Worth, New York, 1982, pp. 105–108.
5. H. Hart, Integration of the rectangular equations of motion in the case of a central force varying inversely as the square of the distance, *Messenger of Math.* 9 (1880) 131–132.
6. T. Heath, *A History of Greek Mathematics*, Vol. 2, Dover, New York, 1981.
7. W. Smart, *Textbook on Spherical Astronomy*, revised by R. Green, Cambridge Univ. Press, 1977. Smart's proof of planar orbits is in the same spirit as the proofs presented above. For instance, Smart does not use vector products.
8. B. Temple and C. Tracy, From Newton to Einstein, *Amer. Math. Monthly*, 99 (1992) 507–521.
9. G. Thomas and R. Finney, *Calculus and Analytic Geometry*, Addison-Wesley, Reading, Massachusetts, Sixth Edition, 1984.
10. R. Weinstock, Inverse-square orbits: Three little-known solutions and a novel integration technique, *Am. J. Phys.*, 60 (1992) 615–619.
11. A. Wintner, *The Analytical Foundations of Celestial Mechanics*, Princeton Univ. Press, 1941. I thank Dr. Robert Weinstock of Oberlin College for telling me of this book, and for his many helpful comments. Dr. Weinstock notes that Wintner's own references are incorrect (see ref. 20 of [10]).

*Northwestern School of Law
Lewis and Clark College
Portland, Oregon 97219*

A Short Proof of a Result on Polynomials

Răzvan Gelca

In this note we want to present a short proof of a result that appeared in [1]. For a polynomial $f(x) = \prod_1^n (x - x_i)$, with distinct real roots $x_1 < x_2 < \cdots < x_n$, we let $d = \delta(f) = \min_i (x_{i+1} - x_i)$ and $g(x) = f'(x)/f(x) = \sum_1^n 1/(x - x_i)$. If k is a real number then the roots of the polynomial $f' - kf$ are also real and distinct.

Proposition. *If for some j , y_0 and y_1 satisfy $y_0 < x_j < y_1 \leq y_0 + d$ then y_0 and y_1 are not zeros of f and $g(y_0) < g(y_1)$.*

Proof: The hypothesis implies that for all i , $y_1 - y_0 \leq d \leq x_{i+1} - x_i$. Hence for $1 \leq i \leq j - 1$ we have $y_0 - x_i \geq y_1 - x_{i+1} > 0$ and so $1/(y_0 - x_i) \leq 1/(y_1 - x_{i+1})$; similarly for $j \leq i \leq n - 1$ we have $y_1 - x_{i+1} \leq y_0 - x_i < 0$ and again $1/(y_0 - x_i) \leq 1/(y_1 - x_{i+1})$.

Finally $y_0 - x_n < 0 < y_1 - x_1$, so $1/(y_0 - x_n) < 0 < 1/(y_1 - x_1)$, and the result follows by addition of these inequalities.

Corollary. $\delta(f' - kf) > \delta(f)$.

Proof: If y_0 and y_1 are zeros of $f' - kf$ with $y_0 < y_1$ then they are separated by a zero of f and satisfy $g(y_0) = g(y_1) = k$. Hence from the proposition we can not have $y_1 \leq y_0 + d$, so $y_1 - y_0 > d$ as required.

REFERENCE

1. P. Walker, Separation of the zeros of polynomials, *Amer. Math. Monthly* (100)(1993) 272–273.

*Department of Mathematics
The University of Iowa
Iowa City, IA 52242
rgelca@math.uiowa.edu*

*Institute of Mathematics
of the Romanian Academy
P.O. Box 1-764
70700 Bucharest Romania*

Two Amusing Dynkin Diagram Graph Classifications

Robert A. Proctor

Here are a couple of simply stated graph classifications which can be used to amuse and amaze students and friends during tea or cocktail parties. It's fun to watch non-mathematicians theologically wrestle with the following notion: Mathematicians can *prove* that no one can come up with any solutions beyond the ones shown in the figures. Many people have been aware of the first classification for some time. The second one is an immediate consequence of a well known fact, but perhaps has not been formulated in this way before.

A *simple graph* is a graph which has no loops or multiple edges. I'll call it *labelled* if a positive real number has been assigned to each vertex.

Problem 1. *Find all connected labelled simple graphs whose labels satisfy the following condition: Twice any label is equal to the sum of the labels of the adjacent vertices.*

Answer. If you check this condition nine times, you can verify that the labels of the last graph in FIGURE 1 satisfy this requirement. For example, at the central vertex we have: $2 \times 6 = 4 + 5 + 3$. Surprisingly, up to an overall scalar multiple of the labels, *all* possible connected graphs labelled in this way are shown in FIGURE 1! There are two infinite families of solutions and then three specific peculiar "exceptional" solutions.

Problem 2. *Find all connected labelled simple graphs whose labels satisfy the following condition: Twice any label minus two is equal to the sum of the labels of the adjacent vertices.*

Answer. The *only* possibilities are shown in FIGURE 2. At the central vertex of the last graph we have $2 \times 270 - 2 = 182 + 220 + 136$. Again there are two infinite families of solutions followed by three exceptional solutions.

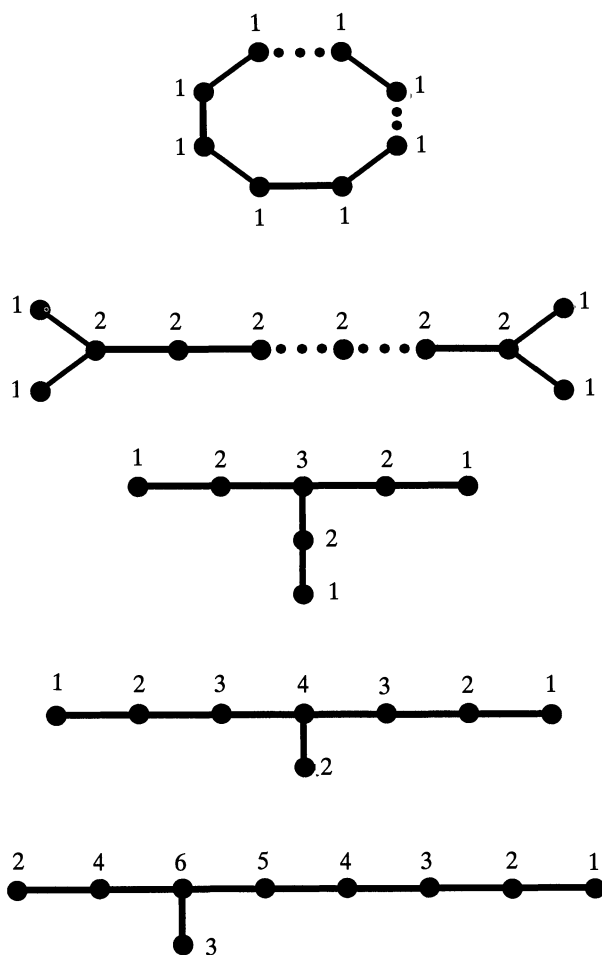
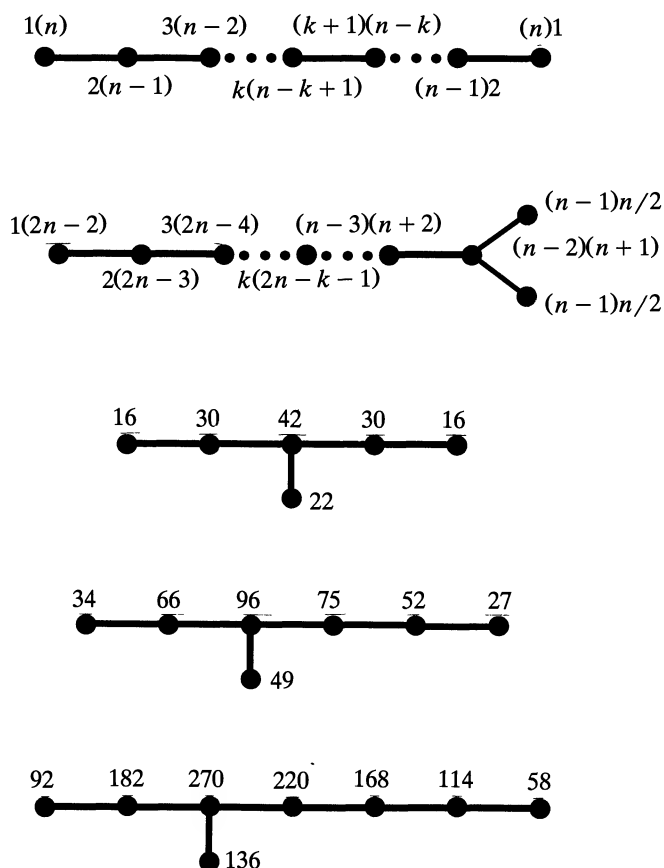


Figure 1

I've used Problem 1 to intrigue friends and students for years, but only recently did I notice Problem 2. I like it better than Problem 1 because the labels are much more entertaining, and because it's easier to explain the significance of its graphs to beginning graduate students: Without the labels, the graphs shown in FIGURE 1 are the extended Dynkin diagrams of types ADE, whereas the graphs of FIGURE 2 without their labels are just the ordinary Dynkin diagrams of types ADE. These play a role in the classification of simple Lie algebras (or groups), whereas the extended diagrams are used to help classify a more sophisticated family of objects, the affine Lie algebras. (Also, the solutions to Problem 2 are unique immediately, without the "overall scalar multiple" fine print needed with the solutions to Problem 1.)

There are many kinds of algebraic and geometric structures arising in mathematics which are "classified" by a list of some kind of Dynkin diagrams. For example, one could ask what are the possible finite subgroups of the orthogonal groups $O(n, \mathbb{R})$ which are generated by reflections. If we ignore the dihedral groups and require that the subgroup fix only the origin, then there is exactly one such subgroup for each member of the following list: $A_n (n \geq 1)$, $B_n (n \geq 2)$, D_n



$(n \geq 4)$, E_6 , E_7 , E_8 , F_4 , G_2 , I_3 , and I_4 . Here each X_n denotes a particular “Coxeter diagram” which has n nodes and which describes a particular subgroup of $O(n, \mathbb{R})$ up to conjugacy according to a certain recipe. These diagrams (shown on page 57 of [BG]) look similar to those appearing in our figures, but the edges are labelled and the vertices are unlabelled. For the details of this classification consult [BG], which was written for undergraduates. The names of our diagrams in FIGURE 1 are $A_n^{(1)}$, $D_n^{(1)}$, $E_6^{(1)}$, $E_7^{(1)}$, and $E_8^{(1)}$ and in FIGURE 2 are A_n , D_n , E_6 , E_7 , and E_8 . Although some aspects of the diagrams and the membership of the list can vary, it is usually readily apparent when a classification by Dynkin diagram-like objects is occurring. The diagrams of type E look quite distinctive, and it seems that one always has at least two diagrams of this type arising. The set of possible simple Lie algebras over \mathbb{C} is indexed by the Dynkin diagrams A_n ($n \geq 1$), B_n ($n \geq 2$), C_n ($n \geq 3$), D_n ($n \geq 4$), E_6 , E_7 , E_8 , F_4 , and G_2 . These diagrams (shown on page 58 of [Hm1]) are exactly what is meant by “Dynkin diagram.” They are similar to Coxeter diagrams, except that some of the edges are directed. Usually the structures of the objects indexed by the version of Dynkin diagram at hand of type A, type D, or type E are easier to deal with than the structures of the objects indexed by the diagrams of other types. This fact is reflected at the diagram level by the diagrams of types ADE having all “easy” edges. The overall phenomenon of classification by Dynkin-like diagrams has many fascinating and mysterious aspects; consult [HHSV] for a survey.

Although the answers to Problems 1 and 2 are stated in terms of graphs, they are actually theorems in linear algebra! To see this, do the following: Let the variables x_1, x_2, \dots denote the as yet unknown vertex labels. In either problem, associate to each labelled graph with n vertices an $n \times n$ system of linear equations. For example, the first two requirements of Problem 2 corresponding to the first two vertices of the first graph in FIGURE 2 give rise to the equations $2x_1 - x_2 = 2$ and $-x_1 + 2x_2 - x_3 = 2$. For the graph of this form with 4 vertices, the system of equations giving all of the requirements for that graph is:

$$\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}.$$

In general, associate to any graph arising in either problem a matrix $A = (a_{ij})$ of the following form: The main diagonal entries $a_{ii} = +2$; the off-diagonal entries a_{ij} are -1 if vertex i is connected to vertex j and 0 otherwise. Conversely, given any $n \times n$ matrix A with $a_{ii} = +2$ and $a_{ij} = a_{ji} = -1$ for some pairs (i, j) and $a_{ij} = 0$ otherwise, one could depict it with a simple graph wherein vertex i is connected to vertex j whenever $a_{ij} = -1$. We say that such a matrix A is *connected* if its corresponding graph is connected. Define three column vectors of length n as follows: $v := (x_1, x_2, \dots, x_n)^T$, $0 := (0, 0, \dots, 0)^T$, and $2 := (2, 2, \dots, 2)^T$. Let's say that v is *positive* if $x_i > 0$ for $1 \leq i \leq n$.

So Problem 1 (respectively Problem 2) actually is asking us to find all connected matrices A of this form for which the linear system $Av = 0$ (respectively $Av = 2$) has a positive solution. With these formulations the answers stated at the beginning of this note are mostly derived on pages 47–54 of [Kac]. These eight pages can be read and understood by themselves, provided that you have had a good course in linear algebra. Kac is interested in such questions because the matrices A , known as generalized Cartan matrices in Lie theory, describe the structure of certain kinds of Lie algebras. On these pages he uses basic linear algebra techniques to investigate the existence of positive solutions v to systems of linear inequalities such as $Av > 0$ and $Av \geq 0$, assuming that A has a certain form. The notions of positive definiteness and positive semi-definiteness play a key role.

Here are the details. Part (e) of Proposition 4.7 and parts (b) and (c) of Theorem 4.8 of [Kac] give the answer to Problem 1. For Problem 2 start with Theorem 4.3. By this result, since A cannot be of type (Aff) or (Ind), it must be of type (Fin). Then part (a) of Theorem 4.8 tells us that the graph S must be one of the graphs listed in FIGURE 2. My contribution is to supply the particular right hand side $(2, 2, \dots, 2)^T$, thereby forming Problem 2 as stated above. It is easy to check that each of the labellings given in FIGURE 2 meet the requirements. In each case only one such labelling is possible, since Theorem 4.3 of [Kac] tells us that $\det A \neq 0$ whenever A is of type (Fin).

The operator that doubles a vertex label and then subtracts the adjacent labels may be thought of as a discrete version of $-\Delta$, where Δ is the Laplace operator.

Now a few comments for people who are familiar with simple Lie algebras. The matrices A associated to each of the graphs of FIGURE 2 are just the Cartan matrices for the root system [Hm1] associated to the graph. Multiplying a column vector from the left by A has the following interpretation: You are just converting a column vector of coordinates with respect to the simple root basis to the fundamental weight basis. Since the coordinates of the famous vector $\rho = \delta$ are

$(1, 1, \dots, 1)^T$ with respect to the fundamental weight basis, the labels appearing on the vertices are just the coordinates of the vector 2ρ in the simple root basis. As such, they appear in tables such as those at the end of [Bou]. The extended Dynkin diagrams of FIGURE 1 can be understood in the context of ordinary root systems as follows. If you adjoin a vertex to the Dynkin diagram of a root system which represents the lowest root $-\beta$ as described on page 95 of [Hm2], then in the ADE cases the diagrams of FIGURE 1 will result. The labels on the remaining vertices give the expansion of β with respect to those simple roots.

Why was I thinking about this recently? In 1980 the labels of FIGURE 2 arose in my thesis (which was written under the direction of Richard Stanley). A member of my committee, George Lusztig, asked me if I knew of an existing interpretation of these mysterious positive integers. Last year while flipping through the recent [MPR], the numbers jumped out at me twelve years late: The typography of the tables in [Bou] was such that I hadn't noticed them before. Fortunately, I was passed on my defense nonetheless! (The paper version of that chapter of my thesis [Pro] describes a Dynkin diagram classification of order diagrams of finite partially ordered sets. That result has a very similar flavor to the subject matter of this note.)

REFERENCES

-
- [BG] L. Grove, C. Benson, *Finite Reflection Groups*, 2nd ed., Springer-Verlag, New York, 1985.
 - [Bou] N. Bourbaki, *Groupes et Algebres de Lie*, Chapitres 4, 5, et 6, Hermann, Paris, 1968.
 - [HHSV] M. Hazewinkel, W. Hesselink, D. Siersma, F. Veldkamp, The ubiquity of Coxeter-Dynkin diagrams, *Nieuw Arch. Wisk.* (8) 25 (1977), 257–307.
 - [Hm1] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, New York, 1972.
 - [Hm2] J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge University Press, Cambridge, 1990.
 - [Kac] V. Kac, *Infinite Dimensional Lie Algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990.
 - [MPR] W. G. McKay, J. Patera, D. W. Rand, *Tables of Representations of Simple Lie Algebras Vol. I*, Les Publications CRM, Montreal, 1990.
 - [Pro] R. Proctor, A Dynkin diagram classification theorem arising from a combinatorial problem, *Advances in Math.* 62 (1986), 103–117.

Department of Mathematics
University of North Carolina
Chapel Hill, NC 27599

Professor Wedderburn's request that the Association be represented on the Editorial Staff of the *Annals of Mathematics* by two associate editors was favorably considered. The Trustees authorized President Ford to appoint a committee of three, including himself, with power to select and nominate two associate editors of the *Annals of Mathematics*. President Ford appointed Professors Cairns and Slaught as the other members of this committee. It was understood that the *Annals* volume will be still further enlarged and it was felt that our subvention to the *Annals* is now inadequate. The Trustees, therefore, voted to increase the annual subvention to \$300.

American Mathematical Monthly
 34, (1927) p. 117

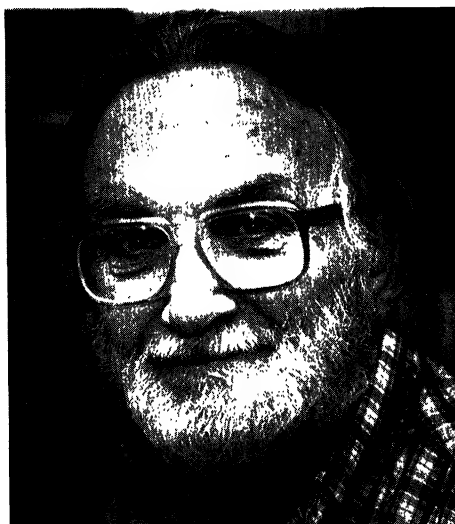
Postcards from Max

As remembered by Paul Halmos

Editor's note: Max Zorn died on March 9, 1993.

Max Zorn was born twenty nine years before Zorn's Lemma, and Zorn's Lemma, the technique and the attitude, will go on living for centuries. For Max the lemma was a remark—the title of his paper on the subject is “A remark on method in transfinite algebra”; it was John Tukey who baptized the result.

Max was a friend of mine, a good friend. We became acquainted in 1969, when I gave a colloquium talk at Bloomington. Max came to the tea before the talk—he came to tea every day, whether there was a colloquium or not—and, in accordance with his custom, he came prepared. On a wrinkled slip of paper (it might actually have been the back of a used envelope) he had scribbled the questions he wanted me to answer—what is my opinion on the work of so-and-so?, how is this work connected with something I wrote ten years before?, has there been any recent progress along the lines of such-and-such? I don't remember any colloquium at which he didn't ask a question afterward (and sometimes during)—a relevant question, a pertinent question, a sharp question. His questions showed that he understood the subject, understood the talk, and was ready to understand and remember the answers. His questions were not intended to be embarrassing, but if the speaker was not thoroughly checked out on all aspects of the subject of his own talk, they could become embarrassing. Max didn't mean to cause pain, and he cheerfully indicated a friendly acceptance of even a vague answer.



Does everybody remember the Piccayune Sentinel? Yes, I spelled it right—the misspelling is Max's own and I faithfully copied both c's. I don't know just when he started it; the first issue that I have a copy of is dated November 1950. It was a one-sheet affair that Max called the world's smallest newspaper and that he gave to a few friends (usually by putting copies into his colleagues' mailboxes, and rarely, for distant friends, by mailing them). One issue I have is labelled “partially late”. The contents of the Piccayune Sentinel were of the same kind as Max himself and his postcards (and as unpredictable and as confusion-inducing)—just longer and more widely distributed.

We were colleagues at Indiana for many years, and we had a routine: most afternoons we would troop over to the physicists' common room in Swain West (the mathematicians couldn't afford such a large and elegant place), get our coffee and cookies, and sit gossiping on the couch by the permanently curtained windows (heaven forbid that some unwanted light or air should enter). Our gossip was never malicious (well, hardly ever): it was about people in the profession (who is moving where and how much will he get paid?), about the profession (could square-summable power series really be relevant to the Riemann hypothesis?), about local matters (who will teach what when and will that room be big enough to hold the class?)—and about books, about movies, about travel, about languages, about anything that had a momentary or a permanent interest for at least one of us. We never ran out of subjects; I looked forward to our meetings, and when some catastrophe prevented one, I missed it.

One conversation we had bothered me afterward, and I was moved to write down my concern in a letter to Max. The letter didn't go through the U.S. Mail—I just put it into Max's box. Here is what I wrote.

"Something you said yesterday worries me—I kept thinking about it during the night and it kept worrying me. You said that you had bad judgement and that you were a failure—two statements with which I thoroughly disagree.

"Of all people I know you are the one who has the sharpest, finest, clearest insight into all of mathematics, and, for that matter, into most of human life. Your tastes and mine (in mathematics, and sometimes in other things) are not always the same—but your insight and your judgement are impressive.

As for failure: that's nonsense. You are respected by everyone who knows you (and by thousands of others), and you are liked by everyone who knows you or ever came anywhere near you. You are a mathematician. Most people no longer know whether your work was algebra, or complex functions, or something funny about semigroups, or whatever—but they respect you for your reputation, for (if you'll pardon the expression) your lemma, for your questions, for your wit (a non-accidental cognate of *Wissenschaft*), for your understanding. You have written, you have taught, you have inspired—is that a failure? I wish I were one!"

Max answered me with a hand-written note that I found in my box the next day.

"In school I heard:

Eigenlob stinkt,
Freundeslob hinkt,
Feindeslob klingt.

(But) thanks."

I wasn't quite sure of all the verbs, so I checked them in a dictionary; roughly (not too roughly) they mean stinks, limps, and rings (respectively).

I left Indiana twice—meaning that I accepted an invitation from another university, moved, returned after a couple of years, and some years later moved again—and we started corresponding. The first time, when at tea one day I said "Max, I'll be leaving", he said "For bad?". That, by the way, is typical of his use of language—he knew idiomatic English perfectly, and had enough control over it that he could twist it to communicate delicate shades of meaning elegantly and efficiently.

He was an unpredictable correspondent. I am a garrulous one—I tend to write repetitive letters full of many details that probably no one besides me is interested in—and he varied from stories with smiles in them to almost brusquely short hello-good-byes. As the years went on, I got in the habit of writing him a longer letter (three or four single-spaced pages) approximately once a month, and he got

in the habit of a short and mysterious friendly postcard approximately twice a year. The operative word is mysterious: he used abbreviations that he invented as he was writing, and with them he referred to happenings, past and future, that I had no way of knowing anything about. Every now and then I really wanted to understand the latest mystery and I demanded an explanation (in my next letter, or even by telephone)—and he was always goodnatured about it, and while seemingly puzzled that someone could fail to understand something that clear, he cheerfully explained. The result was sometimes understandable.

The first Zorn letter that I saved was one of the long friendly kind, two hand-written pages, and it is signed: “as never before, Max”. A few years later another letter ends with: “as before, Max”. One postcard consisted of the following sentence: “If $f(x, y)$ is such that $f(1, y), f(2, y), \dots, f(n, y), \dots$ are computable, then I want $f(x, y)$ to be computable”. A couple of years later (again a postcard): “ \exists I (Nach Kant ist die Existenz des eigenen Ich nicht trivial.)” Still another: “Is the symbol of the symbol (defined and) the same as the symbol?” Again: “Is a random variable a function or an equivalence class of functions?” And: “Sum, ergo dubito.” One letter I received from Max was one typewritten page, on the back of which appeared a backward carbon copy of the same letter, and a handwritten footnote: “You can see that I tried to keep a copy. Long live Freud!”

The letters and postcards came oftener at the beginning than later on—perhaps six to eight times a year—toward the end I was lucky if I got two a year. His last letter came in December 1992; it ends with “I plead fatigue, Max.”

I miss Max.

UNSOLVED PROBLEMS

Edited by: Richard Guy

In this department the MONTHLY presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial or related results. Typescripts should be sent to Richard Guy, Department of Mathematics & Statistics, The University of Calgary, Alberta, Canada T2N 1N4.

A Quarter Century of *Monthly* Unsolved Problems, 1969–1993

Richard K. Guy

A most valuable and timely contribution by Stanley Rabinowitz, which will hopefully greatly reduce the amount of duplication and rediscovery that presently occurs, and will be a boon to all editors of problems sections of all kinds, is his series

Index to Mathematical Problems

of which Volume 1, 1980–1984, is available. It is obtainable from MathPro Press, Westford MA.

References in brackets are to year and page numbers of this MONTHLY, while dates in parentheses refer to publications listed at the end, and other items are labelled (tbp) if they are likely to be published formally, or as written communications (wrc) if publication plans are not presently known. Dates and pages in brackets are also appended to items in the bibliography indicating where the problem originally appeared in the MONTHLY.

In [1969, 54] Victor Klee launched this section of the MONTHLY with the notorious equichordal problem, which goes back to World War I. It gets a mention in reviews by both DeTurck (1993) and Falconer (1993), but even as they wrote the problem was finally being solved by Marek Rychlik (tbp).

The graceful graph [1969, 1128] bibliography is now best regarded as the purview of Joseph Gallian, to whom items should be sent. My somewhat out-of-date version contains 232 papers by 410 authors, only 169 distinct.

Klee [1970, 63] asked for the maximum length of a d -dimensional snake, where by **snake** is meant a simple circuit in the d -cube which has no chords. If we denote this maximum length (number of edges) by $s(d)$, then Abbott and Katchalski (1991) show that $s(d) \geq 77 \times 2^{d-8}$. Their paper contains a very good bibliography.

Erdős and Guy [1973, 52] raised several questions concerning the crossing numbers of graphs; Sýkora and Vrto (1992) give the following lower bounds for the crossing number of the complete bipartite graph, the edge-skeleton of the n -

dimensional cube, and the complete graph, on an orientable surface of genus g .

$$\nu_g(K_{m,n}) > \frac{m^2 n^2}{1200g} - \frac{mn(m+n)}{2} \quad \nu_g(Q_n) > \frac{4^n}{1500g} - n^2 2^{n-1}$$

$$\nu_g(K_n) > \frac{n^4}{6075g} - \frac{n^3}{2}$$

Steven Finch extended Queneau's computations of "Ulam sequences" [1973, 919; 1975, 998; 1987, 962], a (u, v) -**sequence** of positive integers $\{a_i\}$ being defined by $a_1 = u$, $a_2 = v$ and, for $n > 2$, a_n is the least integer expressible *uniquely* as the sum of two distinct earlier members. Queneau showed that the $(2, 5)$ -, $(2, 7)$ - and $(2, 9)$ -sequences are **regular** in the sense that their differences are ultimately periodic. Finch (1991, 1992) proved that if the (u, v) -sequence has only finitely many even terms, then it is regular. Schmerl and Spiegel (tbp) prove that the $(2, v)$ -sequence has just two even terms for any odd $v > 3$.

Leech [1975, 923] asked, for each integer n , what is the greatest integer N such that there exists a tree with n nodes, and edges labelled with integers, in which the distances between pairs of nodes include the consecutive values $1, 2, \dots, N$? Here the distance is the sum of the labels on the unique path joining the nodes. Work of Gibbs and Slater (1991), Herbert Taylor (1991) and Yang Yuan-Sheng (wrc) has improved the results for paths and for more general trees to

n	2	3	4	5	6	7	8	9	10	11	12
paths	1	3	6	9	13	18	24	29	37	45	(51)
trees	1	3	6	9	15	20	26	34	41	(48)	(55)

where the entries in parentheses are not necessarily best possible.

Joseph Gerver (tbp) and evidently Ben Logan before him in 1976, probably found the maximum area sofa that you can move round a corner [1976, 188 and see 1977, 811 and 1991, 974]. A partial description of it is given by Ian Stewart (1992). Its boundary comprises three straight line segments, four arcs of radius $\frac{1}{2}$, seven arcs of involutes of a circle, and four arcs of involutes of involutes of a circle. Its area is ≈ 2.2195 .

In [1983, 35] I warned readers not to try to solve various problems, one of which was the notorious $3x + 1$ problem. Two years later [1985, 3] Lagarias gave a valuable survey and bibliography. Recently he and Weiss (1992) have given two interesting stochastic models for the problem which independently produce the same constant $\gamma_0 \approx 41.677647$ for $\limsup_{n \rightarrow \infty} (\sigma_\infty(n)/\ln n)$, where $\sigma_\infty(n)$ is the number of iterations of the famous function $T(n) = n/2$ (n even), $T(n) = (3n + 1)/2$ (n odd) required to get to the value 1.

In [1991, 974–975] we compared the problem of Forcade, Lamoreaux and Pollington [1986, 119; 1989, 905] with the special case asked by Basil Gordon. The papers of Chandler (1988) and of Forcade and Pollington (1990) are relevant. Blair Kelly III has done a computer search, revealing that $n = 85$ is the smallest counterexample. The next counterexamples are for $92 \leq n \leq 108$, $n = 112$, $n = 113$, $115 \leq n \leq 118$ and $121 \leq n \leq 156$. He says that it is natural to conjecture that there are no Gordon maps for $n > 120$.

Tomaszewski [1986, 280] considered n real numbers a_1, \dots, a_n satisfying $\sum_{i=1}^n a_i^2 = 1$ and asked if, of the 2^n sums of the form $\sum \pm a_i$, it is possible that there are more with $|\sum \pm a_i| > 1$ than there are with $|\sum \pm a_i| \leq 1$. Holzman and Kleitman (tbp) establish the sharp lower bound $3/8$ for the case where the inequality is strict, $|\sum \pm a_i| < 1$, but for the original problem the gap between $3/8$ and the conjectured $1/2$ is still open.

Terry Raines (wrc) says that Erdős, and not your editor, was right: Pambuccian [1986, 627] asked for $a(n)$, the smallest integer a for which there's an integer b , $0 < b < a$, $a \perp b$, such that $a + b, 2a + b, \dots, na + b$ are all composite and asked if $a(n)$ was always prime. Raines notes that for $n = 135$, $a(n) = 8207 = 29 \cdot 283$, with $b = 3251$. He has carried his computations to $n = 180$, and each of $150 \leq n \leq 173$ provide further counterexamples.

In [1988, 927] Tony Gardiner showed that the following four questions are equivalent: for which primes p , if any, (A) is $\binom{2p}{p} \equiv 2 \pmod{p^4}$? (B) is $\sum_1^{p-1} r^{-1} \equiv 0 \pmod{p^3}$? (C) is $\sum_1^{p-1} r^{-2} \equiv 0 \pmod{p^2}$? (D) does p divide the numerator of the Bernoulli number B_{p-3} ? Scott Hochwald (tbp) notes that the questions are indeed equivalent, but that Gardiner's final congruence is incorrect and should be replaced by the statement that

$$S_{p-3}(p) + \frac{p-3}{2} [(p-1)!]^2 \left(1 + \frac{1}{2^2} + \cdots + \frac{1}{(p-1)^2} \right)$$

is divisible by p^2 . The only known prime was 16843, but on a recent visit to Calgary Richard Macintosh found a second example, 2124679.

In [1989, 31] R. J. McG. Dawson asked if there was a subset of a square that contains disjoint connected sets A and B each containing two opposite corners, but does not contain two disjoint connected sets each containing two adjacent corners. Keith Whittington (1991) provides the counterintuitive affirmative answer.

In [1989, 129] Clark Carroll asked for polynomials with integer roots whose derivatives all have integer roots. For cubics the answer is known and can be found, for example, in Walter (1987) or in Buddenhagen, Ford and May (1992); see also MONTHLY problem E3221, solved in [1989, 841–842]. For quartics, there are unpublished papers of Zagier (wrc), Buddenhagen and Ford (wrc) and the present writer, who may have been misleading in [1989, 907–908]. The situation is that for quartics with a repeated root there is an infinity of solutions, given essentially by the rational points on the elliptic curve $y^2 = x^3 - 156x + 560$, **57612** in Cremona (1992). It seems unlikely that there are quartics with all roots distinct, nor higher degree polynomials unless they have sufficiently many repeated roots, but these may still be open questions.

In connexion with Sands's guessing game [1990, 314], see Joel Spencer's (1992) paper.

I apologize that what were offered as unsolved problems in [1992, 74] are in fact well known results. Many of the big names in combinatorial number theory are among those who have written to say that Matiyasevich's generalized harmonic numbers are essentially Stirling numbers of the first kind, and that his conjectures follow fairly easily from known properties. See especially Glaisher (1900), but also Nielsen (1906), Carlitz (1953), Olsen (1966) and Comtet (1974).

In [1992, 178] John Connett asked if a bottle with an inside perfectly reflecting surface could be designed so that a beam of light shone into it was permanently trapped. Robert Dawson, Jan Mycielski and Lior Pachter immediately and independently designed such bottles; their results have been combined (1993). Other solutions were received from M. E. Taylor (wrc), from Madhu Vairy Nayakkankupam (wrc) and from the PCC Rock Creek Math Club—see Bercowitz et al. (wrc).

The page numbers for Connett's second reference should be 1113–1122.

In connexion with the Gordon game [1992, 567] Bob Kibler writes: For the fourteen groups of order 16, White wins only in the cyclic case (by playing to 8). In D_5 White wins by playing to an element of order 5. In D_7 by playing to an element

of order 2. In Z_{12} and Z_{14} by playing to the element of order 2. Black wins in Z_{15} , Z_{17} , D_6 and $Z_2 \times Z_2 \times Z_3$. In Z_{18} White wins by playing to 6—does he also win by playing to 9?

Fatin Sezgin (wrc) applied various tests to the Mycielski sequence [1992, 373] as a result of which he asserts that it cannot be considered as random.

Neil Calkin notes the relevance of Peter Cameron's survey article (1987) and his own thesis (1988) to Steven Finch's 0-additive sequences problem [1992, 671]. Finch has calculated $1\frac{1}{2}$ million terms of the sequence $\{a_n\}$, where $\{a_1, \dots, a_6\} = \{3, 4, 6, 9, 10, 17\}$ and for $n \geq 6$, a_{n+1} is the least integer greater than a_n which is *not* of the form $a_i + a_j$, $i < j$; without detecting any regularity (ultimate periodicity of the differences). Finch believes that this may be due to a massive initial segment of irregular values, while Calkin suspects that there may be counterexamples to Finch's conjecture. They are preparing a joint paper.

David Callan (wrc) solved Parker's permutation problem [1993, 287] affirmatively, and gave an alternative proof that it involves the Catalan numbers. Volker Strehl notes that the problem is not new, and has been solved both qualitatively and quantitatively. It appears, in the 'Griggs' version of the last three lines of [1993, 289], in various contexts: completion of latin squares, a bus scheduling problem, number of terms in the permanent of a circulant matrix. Marshall Hall (1952) attributes the problem to George Cramer, and generalizes and solves it for general finite abelian groups. Marica and Schönheim (1969) apply the result to latin square completion. Brualdi and Newman (1970) solve the enumeration problem by a method closely paralleling that of Gessel in the article under discussion. Chang (1979) uses Hall's theorem and cites Marica and Schönheim. Salzborn and Szekeres (1979) prove Hall's theorem but give no references to earlier work; their motivation was a bus scheduling problem.

REFERENCES

-
- Harvey L. Abbott and M. Katchalski, On the construction of snake in the box codes. *Utilitas Math.*, 40 (1991) 97–116. [1970, 63]
- S. Berkowitz, S. Dahlstrom, G. Madrid, M. McKelvey and F. W. Simmons, Solution to Connett's light-trapping problem (preprint). [1992, 178]
- R. A. Brualdi and M. Newman, An enumeration problem for a congruence equation, *J. Res. Nat. Bur. Stand.*, 74B (1970) 37–40; *MR* 42 #1753. [1993, 287]
- Jim Buddenhagen, Charles Ford and Mike May, Nice cubic polynomials, Pythagorean triples and the law of cosines, *Math. Mag.* 65 (1992) 244–249. [1989, 129]
- Jim Buddenhagen and Charles Ford (wrc), Nice polynomials (1992 preprint). [1989, 129]
- David Callan, Parker's conjecture is true! (preprint, April, 1993) [1993, 287]
- Neil J. Calkin, Sum-free sets and measure spaces, PhD thesis, Univ. of Waterloo, 1988. [1992, 671]
- Neil J. Calkin and Steven R. Finch, Necessary and sufficient conditions for a sum-free set to be ultimately periodic. [1992, 671]
- Peter J. Cameron, Portrait of a typical sum-free set, *Surveys in Combinatorics 1987*, *London Math. Soc. Lecture Notes*, 123 (1987) Cambridge Univ. Press, 13–42. [1992, 671]
- Leonard Carlitz, Note on a theorem of Glaisher, *J. London Math. Soc.*, 28 (1953) 245–246; *MR* 14 726b; rNT B72-12. [1992, 74]
- Leonard Carlitz, A Theorem of Glaisher, *Canad. J. Math.*, 5 (1953) 306–316; *MR* 14 1064b; rNT B72-15. [1992, 74]
- K. A. Chandler, Groups formed by redefining multiplication, *Canad. Math. Bull.*, 31 (1988) 419–423; *MR* 89m: 20021. [1986, 119]
- Gerard J. Chang, Complete diagonals of latin squares, *Canad. Math. Bull.*, 22 (1979) 477–481; *MR* 81g:05033. [1993, 287]
- Louis Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, 1974, p. 229. [1992, 74]
- John E. Cremona, *Algorithms for Modular Elliptic Curves*, Cambridge University Press, 1992.

- Dennis DeTurck, Review of *Unsolved Problems in Geometry*, *Math. Intelligencer*, 15 (1993) 71–72. [1969, 54]
- Kenneth Falconer, Review of *Old and New Unsolved Problems in Plane Geometry and Number Theory*, *Math. Intelligencer*, 15 (1993) 72–74. [1969, 54]
- S. Finch, Conjectures about s -additive sequences, *Fibonacci Quart.*, 29 (1991) 209–214. [1973, 919]
- S. Finch, On the regularity of certain 1-additive sequences, *J. Combin. Theory Ser. A*, 60 (1992) 123–130. [1973, 919]
- S. Finch, Patterns in 1-additive sequences, *Experimental Math.*, 1 (1992) 57–63. [1973, 919]
- R. W. Forcade and A. D. Pollington, What is special about 195? Groups, n th power maps and a problem of Graham, in R. A. Mollin (editor) *Number Theory, Proc. 1st Conf. Canad. Number Theory Assoc., Banff, 1988*, de Gruyter, 1990, 147–155. [1986, 119]
- Joseph L. Gerver, On moving a sofa around a corner, *Geom. Dedicata*, 42 (1992) 267–283; *MR* 93d:51040. [1976, 188]
- Richard A. Gibbs and Peter J. Slater, Distinct distance sets in a graph, *Discrete Math.*, 93 (1991) 155–165. [1975, 923]
- J. W. L. Glaisher, Congruences relating to the sums of products of the first n numbers and to other sums of products, *Quart. J. Pure Appl. Math.*, 31 (1900) 1–35. [1992, 74]
- Marshall Hall, A combinatorial problem on abelian groups, *Proc. Amer. Math. Soc.*, 3 (1952) 584–587; *MR* 14, 350b. [1993, 287]
- Scott H. Hochwald (tbp), How I rediscovered Kummer’s congruence. [1988, 926]
- Ron Holzman and Daniel J. Kleitman, On the product of sign vectors and unit vectors, *Combinatorica*, (to appear). [1986, 280]
- J. C. Lagarias and A. Weiss, The $3x + 1$ problem: two stochastic models, *Ann. Appl. Probab.*, 2 (1992) 229–261. [1983, 35]
- J. Marica and J. Schönheim, Incomplete diagonals of latin squares, *Canad. Math. Bull.*, 12 (1969) 235; *MR* 40 #55. [1993, 287]
- N. Nielsen, *Handbuch der Theorie der Gammafunktion*, 1906; Chelsea, 1966. [1992, 74]
- F. R. Olsen, An extension of a theorem of Nielsen, *Portugal. Math.*, 25 (1966) 63–66; *MR* 35 #4190; *rNT* B72-28. [1992, 74]
- Marek Rychlik, A complete solution of the equichordal point problem of Fujiwara, Blaschke, Rothe and Weizenböck, *Inventiones Math.*, (1993). [1969, 54]
- Franz J. M. Salzborn and G. Szekeres, A problem in combinatorial group theory, *Ars Combin.*, 7 (1979) 3–5; *MR* 81b:05021. [1993, 287]
- James Schmerl and Eugene Spiegel (tbp), The regularity of some 1-additive sequences, *J. Combin. Theory Ser. A*. [1973, 919]
- Fatin Sezgin, On the Mycielski sequence (preprint) Oct 1992. [1992, 373]
- F. Shahrokhi, O. Sýkora, L. A. Székely and I. Vrto, Lower bounds for the crossing number of a graph drawn on a compact 2-manifold, *Proc. 7th Internat. Conf. Graph Theory, Combin., Algorithms, Appl.* (1992). [1973, 52]
- Joel Spencer, Ulam’s searching game with a fixed number of lies, *Theor. Comput. Sci.*, 95 (1992) 307–321. [1990, 314]
- Ian Stewart, Sofa, so good, Chapter 16 of *Another Fine Math You’ve Got Me Into . . .*, W. H. Freeman, 1992. [1976, 188]
- O. Sýkora and I. Vrto, Edge separators for graphs of bounded genus with applications, *Proc. 17th Internat. Workshop Graph Theoretic Concepts Comput. Sci.*, 1991, Springer Lecture Notes in Comput. Sci., 570 (1992) 159–168. [1973, 52]
- Herbert Taylor, A distinct distance set of 9 nodes in a tree of diameter 36, *Discrete Math.*, 93 (1991) 167–168. [1975, 923]
- M. E. Taylor, Ray trapping obstacles in the plane (preprint). [1992, 178]
- Johann Walter, Über ganze rationale Funktionen dritten Grades mit ganzzahligen Koeffizienten, bei denen Nullstellen und Extrema zugleich ganzzahlig sind, *Praxis Math.*, 29 (1987) 489–492. [1989, 127]
- Keith Whittington, Cross the square, this MONTHLY, 98 (1991) 833–834. [1989, 31]
- Don Zagier (wrc), Quartic polynomials all of whose derivatives have integer roots (1989 preprint). [1989, 127]

Department of Mathematics
University of Calgary
Calgary, Alberta
CANADA T2N 1N4

PROBLEMS AND SOLUTIONS

Edited by:
Richard T. Bumby, Fred Kochman and Douglas B. West

Proposed problems should be sent to the MONTHLY PROBLEMS address given on the inside front cover. Please include solutions, relevant references, etc. Three copies are requested.

Solutions of published problems should arrive before May 31, 1994 at the MONTHLY PROBLEMS address given on the inside front cover. Solutions should be typed with double spacing, including the problem number and the solver's name and mailing address. Two copies suffice. A self-addressed postcard or label should be included if an acknowledgment is desired.

*An asterisk (*) after the number of a problem, or part of a problem, indicates that no solution is currently available. Partial solutions will be useful in such cases. Otherwise, the published solution is likely to be based on a solution which is complete and correct. Of course, an elegant partial solution or a method leading to a more general result is always useful and welcome. In addition, references to other appearances of MONTHLY problems or to solutions of these problems in the literature are also solicited.*

PROBLEMS

10346. *Proposed by David Doster, Choate Rosemary Hall, Wallingford, CT.*

Prove that, for all primes p ,

$$\sum_{k=1}^{p-1} \left\lfloor \frac{k^3}{p} \right\rfloor = \frac{(p-2)(p-1)(p+1)}{4}; \quad (A)$$

and

$$\sum_{k=1}^M \left\lfloor \sqrt[3]{kp} \right\rfloor = \frac{(3p-5)(p-2)(p-1)}{4}, \quad (B)$$

where $M = (p-1)(p-2)$.

10347. *Proposed by T. S. Nanjundiah, University of Mysore, Mysore, India.*

For integer $n \geq 1$, define real numbers R_n by

$$R_1 = 1 \quad R_{k+1} = 1 + \frac{k}{R_k} \quad (k \geq 1).$$

Prove that

$$\sqrt{n - \frac{3}{4}} + \frac{1}{2} \leq R_n \leq \sqrt{n + \frac{1}{4}} + \frac{1}{2}$$

for $n \geq 1$.

10348. *Proposed by Jiang Huanxin, student, FuDan University, ShangHai, China.*

Let D, E, F be distinct points on the sides BC, CA , and AB respectively of $\triangle ABC$. Let $\alpha = \angle BDF$, $\beta = \angle FDA$, $\gamma = \angle ADE$, and $\delta = \angle EDC$. If AD, BE , and CF are concurrent and $\alpha/\beta = \delta/\gamma = m$ ($m \neq 1$), prove that $\alpha = \delta$ and $\beta = \gamma$.

10349. *Proposed by Raphael M. Robinson, University of California, Berkeley, CA.*

The hyperbolic plane is tiled with equilateral triangles meeting seven at each vertex. Can the tiles be colored with seven colors in such a way that no two tiles of the same color meet, even at a vertex? (This problem was suggested to the proposer by David Gale.)

10350. *Proposed by Borislav Lazarov, Sofia, Bulgaria.*

Let M be a set of positive integers. Let P_M be set of all primes that divide elements of M , and let L_M be the set of elements of M having no proper divisor in M . Show that P_M finite implies L_M finite.

10351. *Proposed by Leopold Flatto and Jeffrey C. Lagarias, AT & T Bell Laboratories, Murray Hill, NJ.*

Consider the random power series

$$f(t) = \sum_{n=0}^{\infty} \eta_n t^n,$$

where the η_i are drawn independently from $\{-1, 1\}$, with the probability of $\eta_i = 1$ being p for all i .

(a) If $p = 1/2$, show that $f(t)$ has infinitely many zeros in the interval $(0, 1)$ with probability one.

(b) What happens if $p \neq 1/2$?

10352. *Proposed by Yves Nievergelt, Eastern Washington University, Cheney, WA.*

Let U be an open subset of \mathbb{R}^n with smooth boundary ∂U contained in a ball of radius R .

(a) For $n = 3$, show that $\text{Vol}(U) \leq R \cdot \text{Area}(\partial U)/3$.

(b) Generalize to arbitrary dimensions n .

10353. *Proposed by Barry Powell, Kirkland, WA.*

Show that, for any odd prime p , there do not exist non-zero integers, x, y, z satisfying

$$(x, y) = 1 \quad p \nmid xy \quad x^6 + y^6 = z^p.$$

NOTES

Notes: (10347) A weaker version of this appeared as Problem A2 on the 19th Annual William Lowell Putnam Mathematical Competition (November 1958). **(10349)** Since seven triangles meet at each vertex, the angles in the triangles are all $2\pi/7$. The sum of the angles in each triangle is less than π as required in a hyperbolic plane. **(10353)** See P. Ribenboim, *Thirteen lectures on Fermat's Last Theorem*, especially pp. 67–68 where the Jacobi symbols $\left(\frac{Q_p}{Q_l}\right)$ are evaluated, with $Q_p = Q_p(a, b) = (b^p - a^p)/(b - a)$ with a and b odd, relatively prime, and $a \equiv b \pmod{4}$. Other MONTHLY problems dealing with variations of the Fermat equation are E2771 [1979, 308; 1980, 407] and 6558 [1987, 884; 1990, 434].

SOLUTIONS

Periodicity in Multiplicative Groups

6658 [1991, 445]. *Proposed by L. Van Hamme, Free University of Brussels, Belgium.*

Define a sequence of integers by

$$a(0) = 1, \quad a(n) = \sum_{r=0}^{n-1} \binom{n}{r} a(r) \quad \text{for } n \geq 1,$$

so that $\sum_{n=0}^{\infty} a(n)x^n/n! = (2 - e^x)^{-1}$ for $|x| < \log 2$. (This is sequence 1191 in N. J. Sloane's *Handbook of Integer Sequences*, New York, Academic Press, 1973.)

Prove that if p is a prime number and m is an integer not divisible by p , then

$$a(mp^k + s) \equiv a(mp^{k-1} + s) \pmod{p^k}$$

for k a positive integer and s a nonnegative integer.

Solution I by the proposer. Let $\mathbb{R}[X]$ be the set of all real polynomials considered as an \mathbb{R} -vector space and define a linear map

$$\phi: \mathbb{R}[X] \rightarrow \mathbb{Q} \text{ by } \phi(X^n) = a(n) \quad \text{for } n = 0, 1, 2, \dots$$

Apply ϕ to the identity $(X + 1)^n = \sum_r \binom{n}{r} X^r$. Then, for $n \geq 1$,

$$\phi((X + 1)^n) = a(n) + \sum_{r=0}^{n-1} \binom{n}{r} a(r) = 2a(n) = 2\phi(X^n).$$

Hence, for any polynomial $p(X)$,

$$\phi(p(X + 1)) = 2\phi(p(X)) - p(0).$$

Taking for $p(X)$ the polynomial $\binom{x}{r}$, and using the relation $\binom{x+1}{r} = \binom{x}{r} + \binom{x}{r-1}$ for all $r \geq 1$, we get

$$\phi\left(\binom{X}{r}\right) = \phi\left(\binom{X}{r-1}\right) \quad r \geq 1.$$

Thus, ϕ sends $\binom{x}{r}$ to 1 for $r = 0, 1, 2, \dots$. If a polynomial $p(X)$ takes only integer values for $X = 0, 1, 2, \dots$, then $p(X)$ is of the form $p(X) = \sum_r c_r \binom{x}{r}$ with $c_r \in \mathbb{Z}$, and hence $\phi(p(X))$ is an integer. Now apply this observation to the polynomial

$$p(X) = \frac{X^{mp^k+s} - X^{mp^{k-1}+s}}{p^k}.$$

Since $a^{p^k} \equiv a^{p^{k-1}} \pmod{p^k}$ for all integers a , this polynomial is integer-valued; hence

$$\phi(p(X)) = \frac{a(mp^k + s) - a(mp^{k-1} + s)}{p^k}$$

is an integer, as required.

Solution II by Jens Schwaiger, Universität Graz, Graz, Austria. Since

$$\sum_{n=0}^{\infty} a(n) \frac{x^n}{n!} = (2 - e^x)^{-1} = (1 - (e^x - 1))^{-1} = \sum_{m=0}^{\infty} (e^x - 1)^m$$

and since

$$\frac{(e^x - 1)^j}{j!} = \sum_{l=0}^{\infty} S(l, j) \frac{x^l}{l!},$$

where $S(l, j)$ denotes the Stirling number of the second kind given by

$$S(l, j) = \frac{1}{j!} \sum_{i=0}^j (-1)^i \binom{j}{i} (j-i)^l$$

(cf. Louis Comtet, *Advanced Combinatorics*, D. Reidel, 1974, pp. 204–206) we get

$$a(n) = \sum_{j=0}^{\infty} j! S(n, j) = \sum_{j=0}^{N(n)} j! S(n, j)$$

where $N(n)$ is any integer greater than or equal to n .

Putting $n_k = mp^k + s$ and choosing $N(n_k) = N(n_{k-1}) = n_k$, we thus get

$$a(n_k) - a(n_{k-1}) = \sum_{j=0}^{n_k} \sum_{i=0}^j (-1)^i \binom{j}{i} (j-i)^s ((j-i)^{mp^k} - (j-i)^{mp^{k-1}})$$

yielding the desired result as in Solution I.

Editorial comment. The solutions show that the condition that $p \nmid m$ in the statement is not required.

A related use of the operator ϕ of Solution I can be found in Gian-Carlo Rota, "The number of partitions of a set," this MONTHLY, 71 (1964), 498–504.

Solved also by D. Callan, E. Dobrowolski (Canada), O. P. Lossers (The Netherlands), R. Richberg (Germany), and C. Vanden Eynden.

An Absorbing 4-Digit Number

10194 [1992, 161]. *Proposed by Jiro Fukuta, Gifu-ken, Japan.*

(a) For any four-digit number x in base 12, *excluding the eleven numbers with all digits equal*, form the number $A = a_1a_2a_3a_4$ obtained by arranging the four digits in descending order of magnitude. Next form the number $B = a_3a_4a_1a_2$ obtained by exchanging the first two with the last two digits. Put $K(x) = A - B$ and $K^{i+1}(x) = K(K^i(x))$ for $i = 1, 2, \dots$. Prove that $K^i(x) = 4378$ if $i \geq 5$.

(b) Generalize to the base $3 \cdot 2^n$, $n = 0, 1, 2, \dots$.

Solution by Robin J. Chapman, University of Exeter, United Kingdom. When giving a number by digits, we surround it by parentheses, with commas between digits if needed for clarity. Replacing 12 by $b = 3 \cdot 2^n$, we prove that $K^i(x) = (2^n, 2^n - 1, 2^{n+1} - 1, 2^{n+1})$ if $i \geq 2n + 3$ and x does not have equal digits.

We first prove $K(x)$ has the form $(\alpha, \beta, b - 1 - \alpha, b - 1 - \beta)$, with $0 \leq \alpha, \beta < b$. Since $(b - 1 - \alpha, b - 1 - \beta)_b = (b^2 - 1) - (\alpha\beta)_b$, the four-digit form specified equals $(b^2 - 1)[(\alpha\beta)_b + 1]$. By the definition,

$$K(x) = (b^2 - 1)[(a_1a_2)_b - (a_3a_4)_b].$$

Hence we set $(\alpha\beta)_b = (a_1a_2)_b - (a_3a_4)_b - 1$ to complete the claim. This guarantees that $K(x)$ does not have all digits equal if $n > 0$, since that requires $\alpha = b - 1 - \alpha$, but b is even. When $b = 3$, one can have $K(x) = (1111)$, but only when $A = 2210$. The 12 values of x base 3 having this A must also be excluded.

We next prove $K^2(x)$ has the form $(\gamma, \gamma - 1, b - 1 - \gamma, b - \gamma)$, with $1 \leq \gamma < b$, which we call $N(\gamma)$. By the first claim, the digits of $K(x)$ consist of two pairs summing to $b - 1$; let them be $(c_1c_2c_3c_4)$ when put in descending order. Then $K^2(x) = (\alpha', \beta', b - 1 - \alpha', b - 1 - \beta')$, where $(\alpha'\beta')_b = (c_1c_2)_b - (c_3c_4)_b - 1$. This sets $K^2(x) = N(\gamma)$ with $\gamma = c_1 - c_3$.

Now we compute $K(N(\gamma))$ for $1 \leq \gamma < b$. If $\gamma \geq (b + 1)/2$, then in descending order the digits of $N(\gamma)$ are $\gamma, \gamma - 1, b - \gamma, b - \gamma - 1$, from which we compute $K(N(\gamma)) = N(2\gamma - b)$. If $\gamma \leq (b - 1)/2$, then in descending order the digits of $N(\gamma)$ are $b - \gamma, b - \gamma - 1, \gamma, \gamma - 1$, from which we compute $K(N(\gamma)) = N(b - 2\gamma)$. Finally, if $b > 3$ we may have $\gamma = b/2$, in which case the digits of $N(\gamma)$ are $b/2, b/2, b/2 - 1, b/2 - 1$ and $K(N(\gamma)) = N(1)$. Summarizing, we have $K(N(\gamma)) = N(f(\gamma))$, where

$$f(\gamma) = \begin{cases} 1 & \text{if } \gamma = b/2 \\ |2\gamma - b| & \text{if } \gamma \neq b/2. \end{cases}$$

Note that $f(2^n) = 2^n$.

We now claim that $f^{2n+1}(\gamma) = 2^n$ for all γ with $1 \leq \gamma < b$, from which the result follows immediately. This is trivial if $b = 3$, so we may assume $n > 0$. If $\gamma \notin \{2^n, b/2, 2^{n+1}\}$, then $f(\gamma)$ is divisible by more factors of 2 than γ is. We reach $f^r(\gamma) = b/2$ for some $r \leq n - 1$ or $f^r(\gamma) \in \{2^n, 2^{n+1}\}$ for some $r \leq n$. Since $f(2^{n+1}) = f(2^n) = 2^n$, it suffices to show $f^{n+2}(b/2) = 2^n$. This follows by direct computation, since $f(b/2) = 1$ and $f^t(1) = 3 \cdot 2^n - 2^t$ for $1 \leq t \leq n + 1$.

One can give examples where $2n + 3$ iterations are needed. For $n > 0$, let $x = (b/2 + 2, b/2 + 1, 0, 0)$. Then $K^2(x) = N(3)$. Since $f^{2n}(3) = 2^{n+1}$, $K^{2n+2}(x) \neq N(2^n)$. Hence the bound $i \geq 5$ in the statement of part (a) is incorrect; $i \geq 7$ is needed.

Editorial comment. A. Tissier and S. Sagong noted that in bases other than 2 or $3 \cdot 2^n$ this iteration has no fixed point.

Solved also by J. C. Binz (Switzerland), L. Coutry (Egypt), M. Dindos (Slovakia), F. H. Kierstead, Jr., S. Sagong, R. Stong, National Security Agency Problems Group, and the proposer.

Cutting a Parameterized Circle in Half

10198 [1992, 162]. *Proposed by David M. Bloom, Brooklyn College of CUNY, Brooklyn, NY.*

Suppose f is a continuous map of $[0, 1]$ onto a circle. Prove that there exist two closed subintervals of $[0, 1]$ intersecting in at most one point whose images under f are complementary semicircles (i.e., semicircles intersecting only at their endpoints).

Solution by Richard Stong, Rice University, Houston, TX. View the circle as \mathbb{R}/\mathbb{Z} . Since $[0, 1]$ is simply-connected, f lifts to a continuous map $g: [0, 1] \rightarrow \mathbb{R}$. Let a be the maximum value that g attains and let b be a point where $g(b) = a$. Since f is onto, g must attain values arbitrarily near $a - 1$. Therefore, since g is continuous there must be some point e with $g(e) = a - 1$. Assume for definiteness that $e > b$. Let c and d be respectively the smallest and largest values in $[b, e]$ for which $g(x) = a - 1/2$. Then $[b, c]$ and $[d, e]$ are the desired intervals.

Solved also by K. F. Andersen (Canada), D. W. Bailey, W. H. Beckmann, F. Brulois, R. J. Chapman (U.K.), K. S. Kedlaya (student), Y.-H. Kiem (student, Korea), R. Martin (student), A. Müller (France), A. Nijenhuis, N. Passell, B. Richmond, A. Riese, S. T. Stefanov (Bulgaria), E. Suárez (Spain), J. Vogel, T. Zeanah & E. G. Katsoulis, Northern Kentucky University Problem Group, and the proposer. One incorrect solution was received.

Just Below the Graph of $1/(1-x)$

10209 [1992, 266]. *Proposed by Feng Hanqiao, Shaanxi Normal University, Xian, China, and Siu-Ah Ng, University of Hull, Hull, England.*

For each non-negative integer k , define $a_k(n)$ for non-negative integers n by

$$a_k(0) = 1 \quad \text{and} \quad a_k(i+1) = a_k(i) \left(1 + \frac{1}{k} a_k(i) \right) \quad (i \geq 0).$$

Find $\sup_n a_{mn}(n)$ for $m = 1, 2, \dots$

Solution by Reiner Martin (student), University of California, Los Angeles, CA. We will show that

$$\sup_n a_{mn}(n) = \begin{cases} \infty & \text{for } m = 1, \\ \frac{m}{m-1} & \text{for } m > 1. \end{cases}$$

These expressions follow from the inequalities

$$\frac{k}{k-n} - \frac{kn}{(k-n)^3} \leq a_k(n) \leq \frac{k}{k-n}. \quad (1)$$

The right inequality is valid when $k > n \geq 0$; the left inequality requires the

additional condition that $k \geq n + \sqrt{n}$. Given (1), set $k = mn$ to obtain the result for $m > 1$. Since $a_k(n)$ is clearly an increasing function of n for fixed k , the result for $m = 1$ will follow from the fact that the left side is unbounded as a function of n and k with $k > n$.

We prove (1) by induction on n , the case $n = 0$ being trivial. For the right side, if $k > n + 1$, the inductive step is

$$a_k(n+1) = a_k(n) \left(1 + \frac{1}{k} a_k(n) \right) \leq \frac{k}{k-n} \left(1 + \frac{1}{k} \frac{k}{k-n} \right) \leq \frac{k}{k-n-1},$$

where the rightmost inequality follows from $(k-n-1)(k-n+1) \leq (k-n)^2$. Denote the left side of (1) by $f(k, n)$. In order to use

$$f(k, n) \left(1 + \frac{f(k, n)}{k} \right) \leq a_k(n) \left(1 + \frac{a_k(n)}{k} \right) = a_k(n+1)$$

in the inductive step, we demand $f(k, n) \geq 0$, which is guaranteed by $k > n + \sqrt{n}$. This being so, we then wish to show that

$$f(k, n) \left(1 + \frac{f(k, n)}{k} \right) - f(k, n+1)$$

is positive. This is easily done using a computer algebra package. Multiplying this expression by $(k-n)^6(k-n-1)^3$ yields k times an expression which becomes

$$n^2l^3 + nl^5 + 8nl^4 + 17nl^3 + 15nl^2 + 6nl + n + 2l^5 + 9l^4 + 16l^3 + 14l^2 + 6l + 1$$

on substituting $k = n + 1 + l$. Since this is a polynomial with positive coefficients, the result follows.

Editorial comment. By various means, most solvers related $\sup_n a_{mn}$ to $\sum_{n=0}^{\infty} x^n = 1/(1-x)$. Christopher P. Grant and Thomas Kunkle did so by noting that the sequence $a_k(i)$, $0 \leq i < k$ is the approximation to the solution of $y' = y^2$, $y(0) = 1$ on $[0, 1)$ generated by Euler's method with step size $1/k$.

Solved also by R. J. Chapman (U.K.), C. P. Grant, T. Kunkle, O. P. Lossers (The Netherlands), R. Stong, and the proposers.

REVIVALS

Homeomorphisms of Compact Metric Spaces

6612 [1989, 846; 1991, 663]. *Proposed by Ebrahim Salehi, University of Nevada, Las Vegas, NV.*

Suppose X is a compact metric space with metric d , and suppose $T: X \rightarrow X$ is continuous. If

$$\inf_{n \in \mathbb{N}} d(T^n x, T^n y) > 0$$

for each pair x, y of distinct elements of X , prove that T is onto.

Editorial comment. Shortly after the original publication of a solution, David B. Ellis, Ebrahim Salehi, and John Henry Steelman provided counterexamples to the claim, made in that solution, that

$$d'(x, y) = \inf_{n \geq 0} d(T^n x, T^n y)$$

is a metric. For example, if X consists of the three real numbers $0, 1, x$ with $0 < x < 1/2$, using the metric induced from \mathbb{R} , and T interchanges 0 and 1 while fixing x , then $d'(0, 1) = 1 > 2x = d'(0, x) + d'(x, 1)$. Deeper constructions appear to be needed to solve the problem. The following solution is based on the idea of the *enveloping semigroup*. The enveloping semigroup and related notions have proven to be extremely valuable in topological dynamics (see references). The previous argument claimed to work even when T is not assumed continuous. It is still open to decide if the assumption of continuity is required.

Solution by David B. Ellis, Beloit College, Beloit, WI. In order to define our semigroup of functions, we consider the set $X^X = \{f: X \rightarrow X\}$, of all self maps of X . Note that X^X is a semigroup under composition. We give X^X the topology of pointwise convergence, so that

$$f_\alpha \rightarrow f \Leftrightarrow f_\alpha(x) \rightarrow f(x) \quad \text{for every } x \in X.$$

This makes X^X a compact Hausdorff space. By analogy to the *enveloping semigroup* of (X, T) , we form the closure in X^X of the strictly positive iterates of T :

$$\hat{E}(X, T) = \overline{\{T, T^2, \dots, T^n, \dots\}} \subset X^X.$$

Our solution requires two lemmas concerning $\hat{E}(X, T)$. The first lemma is an immediate consequence of the assumption that T is continuous and the fact that we have given X^X the topology of pointwise convergence.

Lemma 1. *Let X be a compact Hausdorff space and $T: X \rightarrow X$ be continuous. Then*

- (a) *the function $L_T: X^X \rightarrow X^X$ defined by $L_T(p) = T \circ p$ is continuous,*
- (b) *the function $R_p: X^X \rightarrow X^X$ defined by $R_p(q) = q \circ p$ is continuous for every $p \in X^X$,*
- (c) *$\hat{E}(X, T)$ is a subsemigroup of X^X .*

Lemma 2. *Let S be a compact Hausdorff space with a semigroup structure in which R_p , defined as in Lemma 1(b), is continuous for every $p \in S$. Then S contains an element u with $u^2 = u$.*

Proof: We use a Zorn's lemma argument. Let

$$\mathcal{M} = \{M \subset S \mid \emptyset \neq M \text{ is closed and } M^2 \subset M\}.$$

Note that $S \in \mathcal{M}$ so \mathcal{M} is nonempty. If $\{M_\alpha \mid \alpha \in A\}$ is a descending chain of elements of \mathcal{M} , then

$$M = \bigcap_{\alpha \in A} M_\alpha \in \mathcal{M}$$

is an infimum. Applying Zorn's lemma we get a minimal nonempty element $N \in \mathcal{M}$. Let $u \in N$. Then $R_u(N) = Nu$ is a compact, hence closed, subset of S . Since $(Nu)(Nu) = (Nu)u \subset Nu \subset N$, it follows that $Nu = N$ because N is

minimal. Now set

$$Q = \{v \in N \mid vu = u\} = R_u^{-1}(\{u\}) \cap N.$$

Q is nonempty because $u \in N = Nu$; Q is closed because R_u is continuous. Moreover $(v_1 v_2)u = v_1(v_2 u) = v_1 u = u$ for any $v_1, v_2 \in Q$; thus $Q^2 = Q$. The minimality of N implies that $Q = N$. In particular $u \in Q$ so that $u^2 = u$.

We now show how the desired result follows from these two lemmas.

Since X is compact, the image of T is closed. Thus it suffices to show that the image of T is dense. To this end, choose $x \in X$ and let U be any open neighborhood of x . We will show that U intersects the image of T .

By the lemmas, we can find an idempotent $u \in \hat{E}(X, T)$. In particular

$$u(u(x)) = u(x).$$

Now u is a limit point of the strictly positive iterates of T in the topology of pointwise convergence. Thus for any neighborhood V of $u(x)$ there exists $n > 0$ such that

$$T^n(u(x)), T^n(x) \in V.$$

The assumption that $\inf d(T^n x, T^n y) > 0$ when $x \neq y$ implies that $x = u(x)$. Taking $V = U$ we have $T^n(x) \in U$, and hence U intersects the image of T .

REFERENCES

1. J. Auslander, *Minimal Flows and their Extensions*, North Holland, Amsterdam, 1988
2. D. Ellis, "What does Topological Dynamics have to do with Algebra?", *preprint*
3. R. Ellis, *Lectures on Topological Dynamics*, Benjamin, New York, 1969

Extraneous Primes

E 3452 [1991, 645]. *Proposed by C. A. Nicol and J. L. Selfridge, University of South Carolina, Columbia, SC.*

If n is an odd integer greater than 3 and ϕ is the Euler function, prove that there exists a prime p such that $p \mid (2^{\phi(n)} - 1)$ but $p \nmid n$.

Editorial comment. Gerry Myerson has pointed out that an extension of the result was misstated, and two values were omitted from what was claimed to be a "complete list". The items listed are the set of pairs a, n such that $(a, n) = 1$, $a > 1$ and $n > 2$ for which there is *no* prime p such that $p \mid (a^{\phi(n)} - 1)$ but $p \nmid n$. The complete list of such (n, a) is $\{(3, 2), (4, 3), (6, 2), (6, 3), (6, 5), (6, 7), (6, 17), (10, 3)\}$.

Source-even Orientations of Graphs

E 3462 [1991, 755; 1993, 594]. *Proposed by J. J. Rotman, University of Illinois at Urbana-Champaign, IL.*

Prove that any connected simple graph with an even number of edges has an orientation (assignment of direction to each edge) such that the number of edges leaving each vertex is even.

Editorial comment. Fred Galvin has pointed out that the word "connected" was omitted from his result for infinite graphs. A correct statement is given below.

Let G be a connected infinite graph and let V_F be the set of vertices of finite degree. Then, for any mapping $p: V_F \rightarrow \{0, 1\}$, there is an orientation of G such that, for every vertex $v \in V_F$, the number of edges leaving v has the same parity as $p(v)$.

If the graph G is allowed to have finite components, it is easy to construct counterexamples. In particular, one can take an infinite number of disjoint copies of the graph consisting of two vertices joined by a single edge, with $p(v) = 0$ for all v .

Collaborating editors: *David F. Appleyard, Paul T. Bateman, Bruce C. Berndt, Duane M. Broline, Barry W. Brunson, Frank S. Cater, Gulbank D. Chakerian, Underwood Dudley, Gerald A. Edgar, Michael A. Filaseta, Ira M. Gessel, Richard A. Gibbs, Jerrold R. Griggs, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Murray Klamkin, Daniel J. Kleitman, Frederick W. Luttman, Frank B. Miles, Richard Pfeifer, Stephen L. Portnoy, J. O. Shallit, John Henry Steelman, Kenneth B. Stolarsky, David E. Tepper, Douglas B. Tyler, Daniel Ullman, and William E. Watkins.*

List of referees and guest editors for 1993: Paul T. Bateman, Gilbert Baumslag, Jozsef Beck, John Brillhart, Ezra Brown, Barry W. Brunson, David Cantor, Bille C. Carlson, Frank S. Cater, Gulbank D. Chakerian, David A. Cox, Dennis DeTurck, John Duncan, Gerald A. Edgar, Noam Elkies, Michael A. Filaseta, Peter C. Fishburn, Dan Flath, Ira M. Gessel, Richard A. Gibbs, Bart Goddard, Sheldon Goldstein, Robert Louis Griess Jr., Branko Grünbaum, Leonid Gurvits, Richard K. Guy, Douglas A. Hensley, John R. Isbell, Mourad E. H. Ismail, Paul Kainen, Geoffrey A. Kandall, Murray S. Klamkin, Janos Komlos, Martin Kruskal, Peter S. Landweber, Solomon Leader, Frederick W. Luttman Jr., Richard N. Lyons, Frank B. Miles, Paul Monsky, Benjamin Muckenhoupt, Ram M. Murty, Roger Nussbaum, Beresford N. Parlett, M. J. Pelling, Richard E. Pfeifer, Robert W. Prielipp, Carl Pomerance, Stanley Rabinowitz, Mizanur Rahman, Doris Schattschneider, Peter Scott, Lawrence A. Shepp, Joseph Silverman, Kenneth B. Stolarsky, David E. Tepper, Jerrold Tunnell, Douglas B. Tyler, Daniel Ullman, William C. Waterhouse, William E. Watkins, Jeffrey R. Weeks, Gregory P. Wene, Douglas West, Cunhui Zhang.

Institute For Advanced Study

In describing the new Institute for Advanced Study at Princeton, Professor Veblen said that a few years ago Mr. Bamberger decided to devote his wealth to some useful purpose and through the influence of Mr. Abraham Flexner decided to devote it to a project for the furtherance of pure scholarship. The plan contemplates a small group of mathematicians who will be free to do scientific work involving no bestowal of degrees, large liberty being allowed to the professors in conducting their activities in the form of seminars or formal lectures or none, as they may wish. It is expected that the students will be beyond the stage of the usual graduate student and that mathematicians will come to the Institute for limited periods of time for the purpose of doing some particular piece of work, for writing a book, etc.

—*American Mathematical Monthly*

40, (1933) p. 128

INDEX TO VOLUME 100, 1993

THE AMERICAN MATHEMATICAL MONTHLY

TITLE INDEX

- 100 Years of *Monthly* Editors, John Ewing, 48
- Aperiodic Chaotic Orbits, S. N. MacEachern and L. M. Berliner, 237
- An Application for the Curiosity $(\log_x N)'$, D. A. Wagstaff, T. A. Norman, and D. M. Campbell, 573
- Are There Only Finitely Many Binomial Coefficients with Positive Deficiency?, R. Guy, 398
- An Axiomatic Approach to the Integral, L. Gillman, 16
- Bisectors of Triangles and Tetrahedra, W. A. Beyer and B. Swartz, 626
- Bricklaying and the Hermite Normal Form, W. J. Gilbert, 242
- Chaotic Motion of a Pendulum with Oscillatory Forcing, S. P. H. and J. B. McLeod, 563
- A Characterization of Inner Product Spaces, N. Falkner, 246
- Chebyshev Polynomials and Regular Polygons, D. Y. Savio and E. R. Suryanarayan, 657
- Counting Critical Points of Real Polynomials in Two Variables, A. Durfee, N. Kronenfeld, H. Munson, J. Roy, and I. Westby, 255
- Data Compression, C. C. McGeoch, 493
- Densest Packings of Congruent Circles in an Equilateral Triangle, H. (J.B.M.) Melissen, 916
- The Equal Area Zones Property, B. Richmond and T. Richmond, 475
- The Evil Twin Strategy for a Football Pool, J. DeStefano, P. Doyle, and J. Laurie Snell, 341
- Famous NonMathematicians, S. G. Buyske, 845
- A Fast Pick-Type Approximation for Areas of H -Polygons, D. Ren, K. Kołodziejczyk, G. Murphy, and J. Reay, 669
- The Fifty-Third William Lowell Mathematical Competition, L. F. Klosinski, G. L. Alexanderson, and L. C. Larson, 758
- From the Post-Markov Theorem Through Decision Problems to Public-Key Cryptography, I. L. Anshel and M. Anshel, 835
- The Fundamental Theorem of Linear Algebra, G. Strang, 848
- Graph Theory and the Game of Sprouts, M. Copper, 478
- How to Make Wavelets, R. S. Strichartz, 539
- Hyperbolic Geometry on a Hyperboloid, W. F. Reynolds, 442
- The Index of a Constrained Critical Point, C. Hassell and E. Rees, 772
- Is There a k -Anisohedral Tile for $k \geq 5$?, J. Berglund, 585
- Isogonal Configurations, T. A. Murdoch, 381
- John Marvin Colaw: and The American Mathematical Monthly, J. D. Maxwell, 117
- John Allen Paulos Replies, J. A. Paulos, 740
- Mathematics for Liberal Arts Students, A. W. Briggs, Jr., 162
- A Matrix Maximum, W. C. Waterhouse, 557
- The Minimal Polynomial of $\cos(2\pi/n)$, W. Watkins and J. Zeitlin, 471
- A mod- n Ackermann Function, or What's So Special About 1969?, J. Froemke and J. W. Grossman, 180
- More on Rectangles Tiled by Rectangles, D. G. Mead and S. K. Stein, 641
- Newton and the Transmutation of Force, T. Needham, 119
- A Note on Diophantine Representations, C. Baxa, 138
- On Some Irrational Decimal Fractions, N. Hegyvári, 779
- On Seeing Progressions of Constant Cross Ratio, R. J. Duffin, 38
- Open Problems in Pattern Avoidance, J. Currie, 790
- Parallel Addition, C. C. McGeoch, 868
- Parker's Permutation Problem Involves the Catalan Numbers, R. Guy, 287
- Partnerships, A. H. Schoenfeld, 926
- Pascal's Matrices, G. S. Call and D. J. Velleman, 372
- Pathological Functions for Newton's Method, G. C. Donovan, A. R. Miller, and T. J. Moreland, 53
- Pick's Theorem, B. Grünbaum and G. C. Shephard, 150
- Polar Area Is the Average of Strip Areas, G. Strang, 250
- The Pompeiu Problem, H. T. Laquer, 461
- Postcards From Max, P. Halmos, 942

- The Principal Axis Theorem over Arbitrary Fields, D. Mornhinweg, D.B. Shapiro, and K. G. Valente, 749
- A Quarter Century of Monthly Unsolved Problems, 1969-1993, R. K. Guy, 945
- A Quicker Convergence to Euler's Constant, D. W. DeTemple, 468
- Quotients of Primes, D. Hobby and D. M. Silberger, 50
- Ramanujan --- For Lowbrows, B. C. Berndt and S. Bhargava, 644
- A Really Trivial Proof of the Lucas-Lehmer Test, J. W. Bruce, 370
- Recurrence of Simple Random Walk in the Plane, T. R. Shore and D. B. Tyler, 144
- Reflections on Rippling Water, M. Mendes France, 743
- The Seventy-Fifth Anniversary Celebration, G. B. Price, 4
- A Simple Proof of the Jordan-Alexander Complement Theorem, A. Dold, 856
- A Simple Proof of Pascal's Hexagon Theorem, J. Van Yzeren, 930
- A Simple Heuristic Proof of Hardy and Littlewood's Conjecture B , M. Rubinstein, 456
- Six, Lies, and Calculators, R. M. Corless, 344
- Small-Group Learning, J. Weissglass, 662
- Squaring the Circle with Holes, H. Rummler, 858
- Symmetries of the Cube and Outer Automorphisms of S_6 , T. A. Fournelle, 377
- Szeged in 1934, E.R. Lorch, 219
- Tarski's High School Identities, S. Burris and S. Lee, 231
- Taxicabs and Sums of Two Cubes, J. H. Silverman, 331
- Thomas Archer Hirst --- I. A Yorkshire Surveyor, J. H. Gardner and R.J. Wilson, 435
- Thomas Archer Hirst --- II. Student Days in Germany, J. H. Gardner and R. J. Wilson, 531
- Thomas Archer Hirst --- III. Göttingen and Berlin, J. H. Gardner and R. J. Wilson, 619
- Thomas Archer Hirst --- V. London in the 1860s, J. H. Gardner and R. J. Wilson, 827
- Thomas Archer Hirst --- IV. Queenwood, France and Italy, J. H. Gardner and R. J. Wilson, 723
- Thomas Archer Hirst --- VI. Years of Decline, J. H. Gardner and R.J. Wilson, 907
- Thoughts on Innumeracy: Mathematics Versus the World?, P. L. Renz, 732
- Two-Year Magazine Subscription Rates, U. Dudley, 34
- The Tyranny of Tests, P. Hilton, 365
- Vandermonde Strikes Again, M. S. Grosof and G. Taiani, 575
- Versatile Coins, I. Szalkai and D. Velleman, 26
- A Visual Explanation of Jensen's Inequality, T. Needham, 768
- What is a Napierian Logarithm?, R. Ayoub, 351
- When Does a Polynomial Over a Finite Field Permute the Elements of Field?, II, R. Lidl and G. L. Mullen, 71
- Yueh-Gin Gung and Dr. Charles Y. Hu Award for Distinguished Service to H. O. Pollak, R. Gnandesikan and H. J. Landau, 115
- Zero-Knowledge Proofs, C. C. McGeoch, 682

AUTHOR INDEX

- Alexanderson, Gerald L. *see Klosinski*
- Anshel, Iris Lee and Michael Anshel, From the Post-Markov Theorem..., 835
- Anshel, Michael *see Anshel, Iris Lee*
- Ayoub, Raymond, What is a Napierian Logarithm?, 351
- Baxa, Christoph, A Note on Diophantine Representations, 138
- Berglund, John, Is There a k -Anisohedral Tile for $k \geq 5$?, 585
- Berliner, L. Mark *see MacEachern*
- Berndt, Bruce C. and S. Bhargava, Ramanujan --- For Lowbrows, 644
- Beyer, W. A. and Blair Swartz, Bisectors of Triangles and Tetrahedra, 626
- Bhargava, S. *see Berndt*
- Briggs Jr., Albert W., Mathematics for Liberal Arts Students, 162
- Bruce, J. W., A Really Trivial Proof of the Lucas-Lehmer Test, 370
- Burris, Stanley and Simon Lee, Tarski's High School Identities, 231
- Buyske, Steven G., Famous Non-Mathematicians, 845
- Call, Gregory S. and Daniel J. Velleman, Pascal's Matrices, 372
- Campbell, Douglas A. *see Wagstaff*
- Copper, Mark, Graph Theory and the Game of Sprouts, 478
- Corless, R. M., Six, Lies, and Calculators, 344
- Currie, James, Open Problems in Pattern Avoidance, 790
- DeStefano, Joseph, Peter Doyle, and J. Laurie Snell, The Evil Twin Strategy for a Football Pool, 341
- DeTemple, Duane W., A Quicker Convergence to Euler's Constant, 468
- Dold, Albrecht, A Simple Proof of the Jordan-Alexander Complement Theorem, 856
- Donovan, George C., Arnold R. Miller, and Timothy J. Moreland, Pathological Functions for Newton's Method, 53
- Doyle, Peter *see DeStefano*
- Dudley, Underwood, Two-Year Magazine Subscription Rates, 34
- Duffin, R. J., On Seeing Progressions of Constant Cross Ratio, 38
- Durfee, Alan, Nathan Kronenfeld, Heidi Munson, Jeff Roy, and Ina Westby, Counting Critical Points..., 255
- Ewing, John, 100 Years of *Monthly* Editors, 48
- Falkner, Neil, A Characterization of Inner Product Spaces, 246
- Fournelle, Thomas A., Symmetries of the Cube and Outer Automorphisms of S_6 , 377
- Froemke, Jon and Jerrold W. Grossman, A mod- n Ackermann Function, or What's So Special About 1969?, 180
- Gardner, J. Helen and Robin J. Wilson, Thomas Archer Hirst --- Mathematician Xtravagant III. Göttingen and Berlin, 619
- Gardner, J. Helen and Robin J. Wilson, Thomas Archer Hirst --- Mathematician Xtravagant V. London in the 1860s, 827
- Gardner, J. Helen and Robin J. Wilson, Thomas Archer Hirst --- Mathematician Xtravagant IV. Queenwood, France and Italy, 723
- Gardner, J. Helen and Robin J. Wilson, Thomas Archer Hirst --- Mathematician Xtravagant II. Student Days in Germany, 531
- Gardner, J. Helen and Robin J. Wilson, Thomas Archer Hirst --- Mathematician Xtravagant I. A Yorkshire Surveyor, 435
- Gardner, J. Helen and Robin J. Wilson, Thomas Archer Hirst --- Mathematician Xtravagant VI. Years of Decline, 907
- Gilbert, William J., Bricklaying and the Hermite Normal Form, 242
- Gillman, Leonard, An Axiomatic Approach to the Integral, 16
- Gnandesikan, R. and Henry L. Landau, Yueh-Gin Gung and Dr. Charles Y. Hu Award to Henry O. Pollak, 115
- Groszof, Miriam Schapiro and Geraldine Taiani, Vandermonde Strikes Again, 575
- Grossman, Jerrold W. *see Froemke*
- Grünbaum, Branko and G. C. Shephard, Pick's Theorem, 150
- Guy, Richard, Are There Only Finitely Many Binomial Coefficients with Positive Deficiency?, 398
- Guy, Richard, Parker's Permutation Problem Involves the Catalan Numbers, 287
- Guy, Richard K., A Quarter Century of Monthly Unsolved Problems, 945
- Halmos, Paul, Postcards From Max, 942
- Hassell, Catherine and Elmer Rees, The Index of a Constrained Critical Point, 772
- Hastings, S. P. and J. B. McLeod, Chaotic Motion of a Pendulum with Oscillatory Forcing, 563
- Hegyvári, Norbert, On Some Irrational Decimal Fractions, 779
- Hilton, Peter, The Tyranny of Tests, 365

- Hobby, David and D. M. Silberger, Quotients of Primes, 50
- Klosinski, Leonard F., Gerald L. Alexanderson, and Loren C. Larson, The Fifty-Third Putnam Competition, 758
- Kołodziejczyk, Krzysztof *see Ren*
- Kronenfeld, Nathan *see Durfee*
- Landau, Henry L. *see Gnandesikan*
- Laquer, H. Turner, The Pompeiu Problem, 461
- Larson, Loren C. *see Klosinski*
- Lee, Simon *see Burris*
- Lidl, Rudolf and Gary L. Mullen, When Does a Polynomial Over a Finite Field Permute the Elements of Field?, II, 71
- Lorch, Edgar R., Szeged in 1934, 219
- MacEachern, Steven N. and L. Mark Berliner, Aperiodic Chaotic Orbits, 237
- Maxwell, John D., John Marvin Colaw: and The American Mathematical Monthly, 117
- McGeoch, Catherine C., Zero-Knowledge Proofs, 682
- McGeoch, Catherine C., Data Compression, 493
- McGeoch, Catherine C., Parallel Addition, 868
- McLeod, J. B. *see Hastings*
- Mead, D. G. and S. K. Stein, More on Rectangles Tiled by Rectangles, 641
- Melissen, (J.B.M.) Hans, Densest Packings of Congruent Circles in an Equilateral Triangle, 916
- Mendés-France, Michel, Reflections on Rippling Water, 743
- Miller, Arnold R. *see Donovan*
- Moreland, Timothy J. *see Donovan*
- Mornhinweg, David, Daniel B. Shapiro, and K. G. Valente, The Principal Axis Theorem over Arbitrary Fields, 749
- Mullen, Gary L. *see Lidl*
- Munson, Heidi *see Durfee*
- Murdoch, Timothy A., Isogonal Configurations, 381
- Murphy, Grattan *see Ren*
- Needham, Tristan, A Visual Explanation of Jensen's Inequality, 768
- Needham, Tristan, Newton and the Transmutation of Force, 119
- Norman, Theodore A. *see Wagstaff*
- Paulos, John Allen, John Allen Paulos Replies, 740
- Price, G. Baley, The Seventy-Fifth Anniversary Celebration, 4
- Reay, John *see Ren*
- Rees, Elmer *see Hassell*
- Ren, Ding, Krzysztof Kołodziejczyk, Grattan Murphy and John Reay, A Fast Pick-Type Approximation for H -Polygons, 669
- Renz, Peter L., Thoughts on Innumeracy: Mathematics Versus the World?, 732
- Reynolds, William F., Hyperbolic Geometry on a Hyperboloid, 442
- Richmond, B. and T. Richmond, The Equal Area Zones Property, 475
- Richmond, T. *see Richmond, B.*
- Roy, Jeff *see Durfee*
- Rubinstein, Michael, A Simple Heuristic Proof of Hardy and Littlewood's Conjecture B, 456
- Rummler, Hansklaus, Squaring the Circle with Holes, 858
- Savio, D. Y. and E. R. Suryanarayan, Chebyshev Polynomials and Reg Polygons, 657
- Schoenfeld, Alan H., Partnerships, 926
- Shapiro, Daniel B. *see Mornhinweg*
- Shephard, G. C. *see Grünbaum*
- Shore, Terence R. and Douglas B. Tyler, Recurrence of Simple Random Walk in the Plane, 144
- Silberger, D. M. *see Hobby*
- Silverman, Joseph H., Taxicabs and Sums of Two Cubes, 331
- Snell, J. Laurie *see DeStefano*
- Stein, S. K. *see Mead*
- Strang, Gilbert, Polar Area Is the Average of Strip Areas, 250
- Strang, Gilbert, The Fundamental Theorem of Linear Algebra, 848
- Strichartz, Robert S., How to Make Wavelets, 539
- Suryanarayan, E. R. *see Savio*
- Swartz, Blair *see Beyer*
- Szalkai, István and Dan Velleman, Versatile Coins, 26
- Taiani, Geraldine *see Groszof*
- Tyler, Douglas B. *see Shore*
- Valente, K. G. *see Mornhinweg*
- Van Yzeren, Jan, A Simple Proof of Pascal's Hexagon Theorem, 930
- Velleman, Dan *see Szalkai*
- Velleman, Daniel J. *see Call*
- Wagstaff, David A., Theodore A. Norman, and Douglas A. Campbell, An Application for the Curiosity $(\log_2 N)'$, 573
- Waterhouse, William C., A Matrix Max, 557
- Watkins, William and Joel Zeitlin, The Minimal Polynomial of $\cos(2\pi/n)$, 471
- Weissglass, Julian, Small-Group Learning, 662
- Westby, Ina *see Durfee*
- Wilson, Robin J. *see Gardner*
- Zeitlin, Joel *see Watkins*

NOTES TITLE INDEX

- A Formula and a Proof of the Infinitude of the Primes, Michael Rubinstein, 388
- A Further Simplification of Dixon's Proof of Cauchy's Integral Theorem, Peter A. Loeb, 680
- A Linear Algebra Approach to Cyclic Extensions in Galois Theory, Evan G. Houston, 64
- A Note on an Identity of Ramanujan, T. S. Nanjundiah, 485
- A Note on Fubini's Theorem, Camille Debieve, 281
- A Short Proof for Romberg Integration, T. von Petersdorff, 783
- A Short Proof of a Result on Polynomials, Răzvan Gelca, 936
- A Short Proof of a Theorem of Erdős and Mordell, André Avez, 60
- A Short Proof of Jacobi's Formula for.. as a Sum of Four Squares, George E. Andrews, Shalosh B. Ekhad, and Doron Zeilberger, 274
- A Simple Example on Non-Sequentialness in Topological Spaces, Heinz König, 674
- A Simple Example of Little Big Set, John K. Williams, 172
- Abelian Forcing Sets, Joseph A. Gallian and Michael Reid, 580
- An Elementary Proof of Hilbert's Inequality, Krzysztof Oleszkiewicz, 276
- An Elementary Proof that the Borromean Rings are Non-Splittable, Ollie Nanyes, 786
- Elementary Proof of the Remez Inequality, Borislav Bojanov, 483
- Embedding Countable Groups in 2-Generator Groups, Fred Galvin, 578
- Generators for the Algebra of Symmetric Polynomials, D. G. Mead, 386
- On an Identity of Daubechies, Doron Zeilberger, 487
- Polynomial Root Dragging, Bruce Anderson, 864
- $PSL_2(\mathbb{Z}) = \mathbb{Z}_2 * \mathbb{Z}_3$, Roger C. Alperin, 385
- Regular Simplices in Spaces of Constant Curvature, Horst Martini, 169
- R_n Contains a Division Ring iff R Does, Ayman Badawi, 679
- Separation of the Zeros of Polynomials, Peter Walker, 272
- Sequences with Large Numbers of Prime Values, Ulrich Abel and Hartmut Siebert, 167
- Simplifying the Proof of Dirichlet's Theorem, Paul Monsky, 861
- The Computer Solves the Three Tower Problem, Arthur Engel, 62
- The Mathematical Relationship Between Kepler's Laws and Newton's Laws, Andrew T. Hyman, 932
- The Secant Method and the Golden Mean, Melvin J. Maron and Robert J. Lopez, 676
- The Symmetry Principle for Möbius Transformations, Louis Brickman, 781
- Two Amusing Dynkin Diagram Graph Classifications, Robert A. Proctor, 937
- Why is P^2 not Embeddable in R^3 ?, Hiroshi Maehara, 862

NOTES AUTHOR INDEX

- Abel, Ulrich and Hartmut Siebert, Sequences with Large Numbers of Prime Values, 167
- Alperin, Roger C., $PSL_2(\mathbb{Z}) = \mathbb{Z}_2 * \mathbb{Z}_3$, 385
- Anderson, Bruce, Polynomial Root Dragging, 864
- Andrews, G. E., S. B. Ekhad, and D. Zeilberger, A Short Proof of Jacobi's Formula for as a Sum of Four Squares, 274
- Avez, André, A Short Proof of a Theorem of Erdős and Mordell, 60
- Badawi, Ayman, R_n Contains a Division Ring iff R Does, 679
- Bojanov, Borislav, Elementary Proof of the Remez Inequality, 483
- Brickman, Louis, The Symmetry Principle for Möbius Transformations, 781
- Debieve, Camille, A Note on Fubini's Theorem, 281
- Ekhad, Shalosh B. *see Andrews*
- Engel, Arthur, The Computer Solves the Three Tower Problem, 62

- Gallian, Joseph A. and Michael Reid, Abelian Forcing Sets, 580
- Galvin, Fred, Embedding Countable Groups in 2-Generator Groups, 578
- Gelca, Răzvan, A Short Proof of a Result on Polynomials, 936
- Houston, Evan G., A Linear Algebra Approach to Cyclic Extensions in Galois Theory, 64
- Hyman, Andrew T., The Mathematical Relationship Between Kepler's Laws and Newton's Laws, 932
- König, Heinz, A Simple Example on Non-Sequentialness in Topological Spaces, 674
- Loeb, Peter A., A Further Simplification of Dixon's Proof of Cauchy's Integral Theorem, 680
- Lopez, Robert J. *see* Maron
- Maehara, Hiroshi, Why is P^2 not Embeddable in R^3 ?, 862
- Maron, Melvin J. and Robert J. Lopez, The Secant Method and the Golden Mean, 676
- Martini, Horst, Regular Simplices in Spaces of Constant Curvature, 169
- Mead, D. G., Generators for the Algebra of Symmetric Polynomials, 386
- Monsky, Paul, Simplifying the Proof of Dirichlet's Theorem, 861
- Nanjundiah, T. S., A Note on an Identity of Ramanujan, 485
- Nanyes, Ollie, An Elementary Proof that the Borromean Rings are Non-Splittable, 786
- Oleszkiewicz, Krzysztof, An Elementary Proof of Hilbert's Inequality, 276
- Proctor, Robert A., Two Amusing Dynkin Diagram Graph Classifications, 937
- Reid, Michael *see* Gallian
- Rubinstein, Michael, A Formula and a Proof of the Infinitude of the Primes, 388
- Siebert, Hartmut *see* Abel
- von Petersdorff, T., A Short Proof for Romberg Integration, 783
- Walker, Peter, Separation of the Zeros of Polynomials, 272
- Williams, John K., A Simple Example of Little Big Set, 172
- Zeilberger, Doron, On an Identity of Daubechies, 487
- Zeilberger, Doron *see* Andrews

REVIEWS BY TITLE

Names of authors are in ordinary type; those of reviewers in capitals.

- A First Course in Chaotic Dynamical Systems*, Robert L. Devaney, PHILIP STRAFFIN, 961
- A First Course in Noncommutative Rings*, T.-Y. Lam, LANCE SMALL, 689
- Ethnomathematics: A Multicultural View of Mathematical Ideas*, Marcia Ascher, JUDITH V. GRABINER, 304
- How to Read and Do Proofs*, Daniel Solow, M. F. JANOWITZ, 197
- Iteration of Rational Functions*, Alan F. Beardon, ROBERT L. DEVANEY, 90
- Journey into Geometries*, Marta Sved, WILLIAM E. FENTON, 411
- Linear Algebra Through Geometry*, Thomas Banchoff and John Wermer, JAMES KUZMANOVICH, 506
- Mathematics in Industrial Problems. Parts 1-4*, edited by Avner Friedman, ELLIS CUMBERBATCH, 597
- Ordinary Differential Equations*, Vladimir I. Arnol'd, FRED BRAUER, 810
- Polyominoes: A Guide to Puzzles and Problems in Tiling*, George Martin, VICTOR G. FESER, 412
- Roads to Geometry*, Edward C. Wallace and Stephen F. West, WILLIAM E. FENTON, 411
- Second Year Calculus*, David M. Bressoud, WILLIAM FARIS, 886

SOLUTIONS

Numbers in boldface refer to problems; those in lightface to pages.

6612*	957	10187	695	E3415	84	E3456	798
6634	294	10189	504	E3428	85	E3457	502
6647	186	10190	505	E3439	188	E3459	593
6656	402	10191	695	E3441	86	E3460	87
6658	953	10193	878	E3442	591	E3462	594
6659	500	10194	955	E3444	189	E3462*	959
6660	296	10197	806	E3446	190	E3463	693
6661	191	10198	956	E3447	191	E3465	800
6662	300	10200	807	E3448	298	E3466	409
6663	592	10202	880	E3449	690	E3467	800
6664	692	10205	881	E3450	300	E3468	801
6665	405	10207	882	E3451	302	E3469	875
6669	407	10209	956	E3452	404	E3471	876
6670	595	10217	595	E3452*	959	E3472	503
6671	694	E3267	292	E3453	404		
6672	803	E3394	77	E3454	194		
10184	877	E3412	81	E3455	691		
						*Revivals	

PROBLEMS PROPOSED

Al-Ahmar, M.	76	Galvin, Fred	<i>see Baloglou</i>
Allison, David	290	Gessel, Ira	689
Alzer, Horst	798	Golomb, Solomon W.	499
Anglesio, Jean	291	Golomb, Michael	797
Arterburn, David	<i>see Axness</i>	Gurarie, David E.	401
Austin, A. Keith	689	Guy, Richard K.	589
Austin, A. Keith	185	Hahn, Liang-shin	185
Axness, Carl	590	Hajja, Mowaffaq	589
Bagby, Richard	873	Holton, Derek A.	291
Baloglou, George	874	Huanxin, Jiang	952
Bang, Seung-Jin	498	Huanxin, Jiang	688
Barnett, Jeffrey A.	401	Hubbard John H.	<i>see Connelly</i>
Bloom, David M.	185	Isaacs, I. Martin	<i>see Williams</i>
Bloom, David M.	874	Johnson, Bruce R.	185
Borwein, David	797	Kedlaya, Kiran K.	796
Borwein, Jonathan	<i>see Borwein, David</i>	Klamkin, Murray S. and A. Liu	75
Borwein, Jonathan M.	76	Knuth, Donald E.	76
Byerly, Robert E.	75	Knuth, Donald E.	400
Cain, George	874	Kotlarski, Ignacy I.	797
Canfield, E. Rodney	499	Kotlarski, Ignacy I.	590
Connelly, Robert	498	Kotlarski, Ignacy I.	401
Doster, David	951	Krafft, O.	499
Ehrhart, E.	874	Lagarias, J. C.	402
Erdős, Paul	291	Lagarias, Jeffrey C.	<i>see Flatto</i>
Erdős, Paul	184	Lazarov, Borislav	952
Flatto, Leopold	952	Letac, Gérard	689
Gagola Jr., Stephen M.	76	Liu, A.	<i>see Klamkin</i>

- Lu, Zhiging *see Cain*
 Luo, Feng and Richard Stong 184
 MacKinnon, Nick 590
 Márquez, Juan Bosco Romero 590
 Mason, Eric H. 401
 Mattics, L. E. 76
 Minkus, Jerome 689
 Morris, Howard 290
 Nanjundiah, T. S. 951
 Nievergelt, Yves 952
 Nowakowski, Richard J. *see Guy*
 Oluyede, Broderick 689
 Palacios, José Luis 400
 Penney, David 688
 Pinsky, Mark A. 291
 Pomerance, Carl *see Penney*
 Pomerance, Carl 796
 Popescu, Călin 185
 Powell, Barry 952
 Richter, R. Bruce 796
 Robinson, Raphael M. 76
 Robinson, Raphael M. 952
 Rosenfeld, Moshe 873
 Rosenfeld, Moshe 291
 Rosset, Shmuel 874
 Rubinstein, Zalman 291
 Rudin, Walter 499
 Sarli, John 797
 Schaefer, M. *see Krafft*
 Schäffke, Reinhard *see Axness*
 Schmidt, Frank 185
 Širáň, Josef *see Richter*
 Stong, Richard *see Luo*
 Vanden Eynden, Charles 873
 Vince, Andrew 589
 Wardlaw, William P. 401
 Wardlaw, William P. 688
 Wardlaw, William P. 590
 Whiteley, Walter *see Connelly*
 Williams, John Calvin 498
 Zha, Hongyuan 499

PROBLEMS SOLVED

- Alvis, Dean 84
 Andersen, Kenneth F. 190
 Athanasiadis, Christos 878
 Brulois, Frédéric 690
 Callan, David 801
 Callan, David 404
 Chapman, Robin J. 405
 Chapman, Robin J. 409
 Chapman, Robin J. 877
 Chapman, Robin J. 403
 Chapman, Robin J. 696
 Chapman, Robin J. 955
 Chapman, Robin J. 880
 Chapman, Robin J. 595
 Chen, William Y. C. 800
 Conolly, B. W. 803
 Cruz-Uribe, David 195
 Darling, Donald A. 406
 Ellis, David B. 958
 Erdős, Paul 407
 Flanigan, F. J. 403
 Ford, Kevin 192
 Ford, Kevin 801
 Galvin, Fred 593
 Griffin, Peter 81
 Griggs, Jerrold R. 503
 Grossman, Jerrold 594
 Herman, Eugene A. 86
 Hernández, Victor 87
 Hesterberg, Tim *see Griffin*
 Holzsager, Richard 594
 Holzsager, Richard 693
 Holzsager, Richard 503
 Honold, Thomas 595
 Kanetkar, Sharad 187
 Kastanas, Ilias 693
 Kastanas, Ilias 405
 Kedlaya, Kiran S. 691
 Kedlaya, Kiran S. 188
 Kedlaya, Kiran S. 877
 Komanda, Nasha 300
 Kunkle, Thomas 799
 Lakshmanan, Neela 876
 Laquer, H. Turner 86
 Lindsey II, John H. 693
 Lossers, O. P. 186
 Lossers, O. P. 591
 Lossers, O. P. 78
 Lossers, O. P. 806
 Lossers, O. P. 295
 Lossers, O. P. 191
 Magagnosc, David 303
 Martin, Reiner 956
 Mattics, L. E. 694
 Maus, Sonja *see Honold*
 Nieto, José Heber 299
 Nijenhuis, Albert 876
 Nijenhuis, Albert 292
 Nijenhuis, Albert 877
 Nijenhuis, Albert 87

Nylen, Peter 881
 Ott, Steve 189
 Prasolov, Victor 875
 Richberg, Rolf 188
 Richberg, Rolf 501
 Richberg, Rolf 592
 Rogers, Kenneth 404
 Roth, M. 692
 Schilling, Kenneth 301
 Schilling, Kenneth 504
 Schwaiger, Jens 954
 Shirali, Shailesh 880
 Singer, Nicholas C. 695
 Steelman, John Henry 505
 Stong, Richard 956

Stong, Richard *see Griffin*
 Stong, Richard 878
 Stong, Richard 807
 Šuch, O. *see Roth*
 Tam, Tin-Yau *see Nylen*
 Thoo, J. B. 690
 Tschiersch, Rainer 296
 Uhlig, Frank *see Nylen*
 Van Hamme, L. 953
 Vanden Eynden, Charles 693
 Velleman, Dan 85
 Wengenroth, J. 301
 Wildhagen, Chris *see Richberg*
 Zacharias, Klaus 194

Thanks

The Monthly expresses its appreciation to the following people for their help in refereeing during the past year. We could not function without such people and their hard work.

David J. Aldous, Richard A. Askey, Edward J. Barbeau Jr., John Kelly Beem, Jeffrey M Bergen, Bruce Berndt, T.S. Blyth, Ken Brakke, Richard A. Brualdi, Dorothy D. Buerk, Joe Buhler, William Calhoun, Paul J. Campbell, Timothy J. Carlson, Karen L. Collins, Robert M. Corless, William H. Cunningham, Alejandro de Acosta, Robert Devaney, Edsger W. Dijkstra, Josef F. Dorfmeister, Paul Edelman, Gerald Edgar, Anthony Edwards, James F. Epperson, Clint Erb, Ron Evans, J. Douglas Faires, Lawrence A. Fialkow, Daniel Flath, Ira M. Gessel, Leonard Gilman, John Gimbel, Andrew M. Gleason, John Robert Greene, David Gries, Branko Grünbaum, William H. Gustafson, Wolfgang Haken, Philip J. Hanlon, John G. Harvey, Bruce Hill, Thomas W. Hungerford, Glenn Hurlbert, Joe Iaia, Victor Klee, Ronald Knill, Jeffrey C. Lagarias, Tsit-Yuen Lam, David J. Leeming, James MacQueen, Peter R. Massopust, Stephen B. Maurer, Albert Nijenhuis, Alec Norton, Joseph O'Rourke, Frank Okoh, Ingram Olkin, Richard Olshen, Guillermo Owen, Sharon L. Pedersen, Subramanian Ramakrishnan, Bruce Reznick, Ira Rosenholtz, Cecil C. Rousseau, Patrick J. Ryan, Marco Scarsini, Frederick Jr. Schmitt, Jeffrey Shallit, Daniel B. Shapiro, Stuart J. Sidney, Rodica Simion, Joseph H. Sliverman, Tara L. Smith, J. Laurie Snell, Hunter Snevily, G. Sparling, Joel Spencer, Ralph G. Stanton, Robert Steinberg, Lajos F. Takacs, Alan D. Taylor, Jeremy T. Teitelbaum, Alfred Jacobus Van der Poorten, David C. Vella, William Waterhouse, William Watkins, Jet Wimp, Brian J. Winkel

We give a special thanks to the Notes consultants:

John Akeroyd, Dennis Brewer, Dan Luecking, Mihalís Maliakas, Itrel Monroe, Serge Tabachnikov.